AD A091912

Report NAVTRAEQUIPCEN 79-C-0042-1

LEVEL II

(12)

PRECISION APPROACH RADAR TRAINING SYSTEM (PARTS) TRAINING EFFECTIVENESS EVALUATION

Michael E. McCauley
Clarence A. Semple
Canyon Research Group, Inc.
741 Lakefield Rd., Suite B
Westlake Village, CA 91361

DTIC
NOV 18 1980
C

August 1980

FINAL REPORT FOR PERIOD
23 AUGUST 1979 - 23 AUGUST 1980

DoD Distribution Statement

Approved for public release;
distribution unlimited.

DCASMA
Los Angeles Region, Van Nuys Branch
6230 Van Nuys Blvd.
Van Nuys, CA 91408

NAVAL TRAINING EQUIPMENT CENTER
ORLANDO, FLORIDA 32813

DDC FILE

80 11 14 096

GOVERNMENT RIGHTS IN DATA STATEMENT

Reproduction of this publication in whole or in part is permitted for any purpose of the United States Government.

UNCLASSIFIED 19

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

14 30-3067

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 18 NAVTRAEQUIPC[redacted] 79-C-0042-1
2. GOVT ACCESSION NO.: AD-A091912
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): 6 Precision Approach Radar Training System (PARTS) Training Effectiveness Evaluation
5. TYPE OF REPORT & PERIOD COVERED: 9 Final Report. 23 Aug 79 — 23 Aug 80
6. PERFORMING ORG. REPORT NUMBER: 30-3067
7. AUTHOR(s): 10 Michael E. McCauley [redacted] Clarence A. Semple
8. CONTRACT OR GRANT NUMBER(s): 15 N61339-79-C-0042
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Canyon Research Group, Inc., 741 Lakefield Rd., Suite B, Westlake Village, CA 91361
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NAVTRAEQUIPCEN Task No. 7796-3P1
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Training Equipment Center, Orlando, FL 32813
12. REPORT DATE: 11 Aug[redacted]80
13. NUMBER OF PAGES: 128
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): DCASMA, Los Angeles Region, Van Nuys Branch, 6230 Van Nuys Blvd., Van Nuys, CA 91408 — 12 128
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report):
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report): Approved for public release; distribution unlimited
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): VTAG, Precision Approach Radar, Automated Training, Automated Speech Technology, Computer Speech Recognition, Computer Speech Synthesis, Performance Measurement, Training Effectiveness Evaluation, Transfer of Training, Air Traffic Control
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

The use of automated speech technology (AST) in training systems was addressed through the evaluation of the experimental prototype Precision Approach Radar Training System (PARTS). The use of computer speech recognition for simulating the approach control task and for providing input to the automated performance measurement subsystem demonstrated the potential for enhancing automated training. Computer speech synthesis provided the capability to simulate verbal communications from pilots and other controllers, → next

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 391185

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Continued

and to demonstrate proper precision approach radar (PAR) procedures by a simulated "model" controller.

The PARTS was evaluated during an eight-month period while being used by students at the Naval Air Traffic Control Schools. Students trained on PARTS were compared with students from the normal PAR training course in a pseudo transfer of training study. No significant differences in PAR performance were found, but the entire range of skills taught by PARTS was not able to be included in the criterion test.

Observation, interview, and a performance measurement validation study revealed a number of problems associated with the courseware in PARTS. These problems led to limited acceptance by instructors, but student acceptance was high.

A cost-effectiveness analysis generated estimates of potential savings in personnel utilization through the use of AST training systems.

Suggestions were made for PARTS design modifications and for future applications of AST training systems, particularly in air traffic control.

FOREWORD

Two names have been used to refer to the first application of computer speech recognition technology to a real-world training problem. GCA-CTS (Ground Controlled Approach Controller Training System) is the name Logicon gave to the system they built to provide automated adaptive training for precision approach radar controller training. PARTS (Precision Approach Radar Training System) is the name of the evaluation of the GCA-CTS which was conducted by Canyon Research. Canyon's independent evaluation of GCA-CTS is the topic of this report, which provides guidelines for the modifications required in order to transition the technology of the experimental prototype into an operational training system such as the Navy's device 15G19 for Air Traffic Control radar training. This report clearly states that the prototype, as delivered by Logicon, would require revision of its courseware prior to its use as an operational training system, and just as clearly states what revisions are needed. Thus, this report should be used in conjunction with the Logicon authored NAVTRAEQUIPCEN 77-C-0162 series of technical reports to develop a specification for the operational application of the instructional features embodied in GCA-CTS to Air Traffic Controller radar training devices.

R. BREAUX, PH.D.
Scientific Officer

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A

ACKNOWLEDGEMENTS

This project evaluated the use of automated speech technology for training, as implemented in an experimental prototype system called the Precision Approach Radar Training System (PARTS). The study was done under contract N61339-79-C-0042 from the Naval Training Equipment Center, with Dr. Robert Breaux serving as the Scientific Officer. The period of performance was August 1979 to August 1980.

The authors are grateful to Dr. Breaux for his support and encouragement during this project.

Thanks are extended to the instructors and students who participated in this study at the Air Traffic Control Schools Department, Naval Air Technical Training Center, NAS Memphis. CDR Windsor, Officer in Charge, and his staff provided support and cooperation throughout the project. Thanks are extended to CWO Robert R. Lynchard for his efficient job as ATC Schools coordinator for the study.

The authors are especially grateful for the effort, cooperation, hard work and interest demonstrated by ACC Gerald H. Cyr and AC1 Terry M. Martin. Much of the information in this report originated through their efforts. Personnel from the ATC Schools, however, should not be held responsible for the content, opinions and conclusions expressed in this report, as they are solely the responsibility of the authors.

Thanks to Mary Hicklin, David Harry and Michael Grady of Logicon, Inc. for providing information and documentation in support of the study.

We appreciate the tireless efforts of Rosemary Wescott in producing the project reports.

The authors are indebted to Joan A. Helms and Margaret A. Covington for their diligent work in data collection and analysis.

TABLE OF CONTENTS

TABLE OF CONTENTS
(CONTINUED)

TABLE OF CONTENTS (CONTINUED)

LIST OF ILLUSTRATIONS

LIST OF TABLES

LIST OF TABLES
(CONTINUED)

SECTION I

INTRODUCTION

Recent advances in the technology of computer speech recognition have opened the door to new techniques in man-computer communication. The possible applications of this technology are as unlimited as speech itself.

Automated speech technology and some of its potential applications have been reviewed in several recent publications (Dixon and Martin, 1979; Harris, 1980; Lea 1979, 1980). The U.S. technology has advanced primarily in isolated word recognition (IWR), concentrating on a small vocabulary, discrete words or short phrases, and requiring pre-training by each speaker. This technology only recently has been applied to training systems. The present study reports a field evaluation of a training system incorporating automated speech technology (AST). In this report the use of the term AST will refer to the combination of both computer speech recognition and computer speech generation or synthesis.

The training system evaluated in this study is the experimental prototype Precision Approach Radar Training System (PARTS). It was developed by Logicon Inc., under Naval Training Equipment Center (NAVTRAEQUIPCEN) sponsorship beginning in 1973 and was known during its development as the Ground Controlled Approach - Controller Training System (GCA-CTS).

This study had a dual purpose, as stated in the original study specification from NAVTRAEQUIPCEN: to evaluate the use of automated speech technology for training, and to evaluate, specifically, the training features of PARTS. This dual purpose led to some organizational and conceptual difficulties during conduct of the evaluation, since it is difficult to evaluate a specific example and draw conclusions about a general concept. Nevertheless, Canyon Research Group, Inc. (Canyon) has attempted to address both aspects of the study by evaluating the general conceptions of the use of speech recognition and speech synthesis for training, as well as analyzing the training effectiveness of this specific prototype system.

Specific objectives of the study were to assess the training effectiveness of the prototype PARTS technology, make recommendations to the NAVTRAEQUIPCEN concerning the application of this technology for Precision Approach Radar (PAR) training in the future, and to evaluate a variety of factors which may influence training effectiveness, such as: validity of the performance measurement system; transfer of training; the design of courseware and training features; human factors design of student and instructor stations; and user acceptance. In addition, training efficiency issues were addressed by estimating resource expenditure rates for the traditional training system compared to a hypothetical multi-station PARTS.

This report covers a two-phase, ten-month study which began in August, 1979. In the first phase, many features of the training system were analyzed but the computer speech recognition and automated performance measurement systems were emphasized. In the second phase, emphasis was given to transfer of training (TOT) in which students who were trained on PARTS were compared to traditionally trained students in a TOT performance test. Canyon's general approach to the evaluation is depicted in Figure 1.

Because the objectives of the study were broad, many methods were used to gather information, including review of literature and documentation, experimentation, structured observation, questionnaire survey and interviews. In short, any method was used that could help gain information about the structure, function or training value of PARTS and its advanced technologies.

The analysis and evaluation of the PARTS was made more challenging by its changing nature. The study began at approximately the same time that the prototype PARTS was installed at the Naval Air Technical Training Center (NATTC), NAS Memphis, Millington, Tennessee. There was little time for debugging software prior to using the system with students at the Air Traffic Control Schools. The system was plagued with breakdowns early in the evaluation. The cause of these breakdowns was a combination of hardware, software and local power fluctuations ranging from less than 100 to nearly 250 volts. These initial adjustment problems seemed to be remedied by the end of the evaluation period, but changes in the system occurred frequently during that time, as is the nature of an experimental prototype system. Evaluation of the system was similar to a tracking problem -- locating a moving target. However, none of these challenging issues was totally unexpected in a field evaluation and the analyses, conclusions and recommendations of this report are based on the more enduring characteristics of the system rather than the transient ones.

To preview this report, first a discussion of the use of automated speech technology in training systems is given. Then the background and development of the experimental prototype PARTS is be discussed, followed by a discussion of speech recognition and synthesis in PARTS. Subsequent topics include the transfer of training study, factors affecting the training effectiveness of PARTS, user acceptance, training efficiency (cost implications), and conclusions and recommendations. The reader who is interested in general issues of AST, rather than specific analyses of the prototype PARTS, would do well to read Sections II, X, and XI. Emphasis on PARTS is given in Sections III through IX.

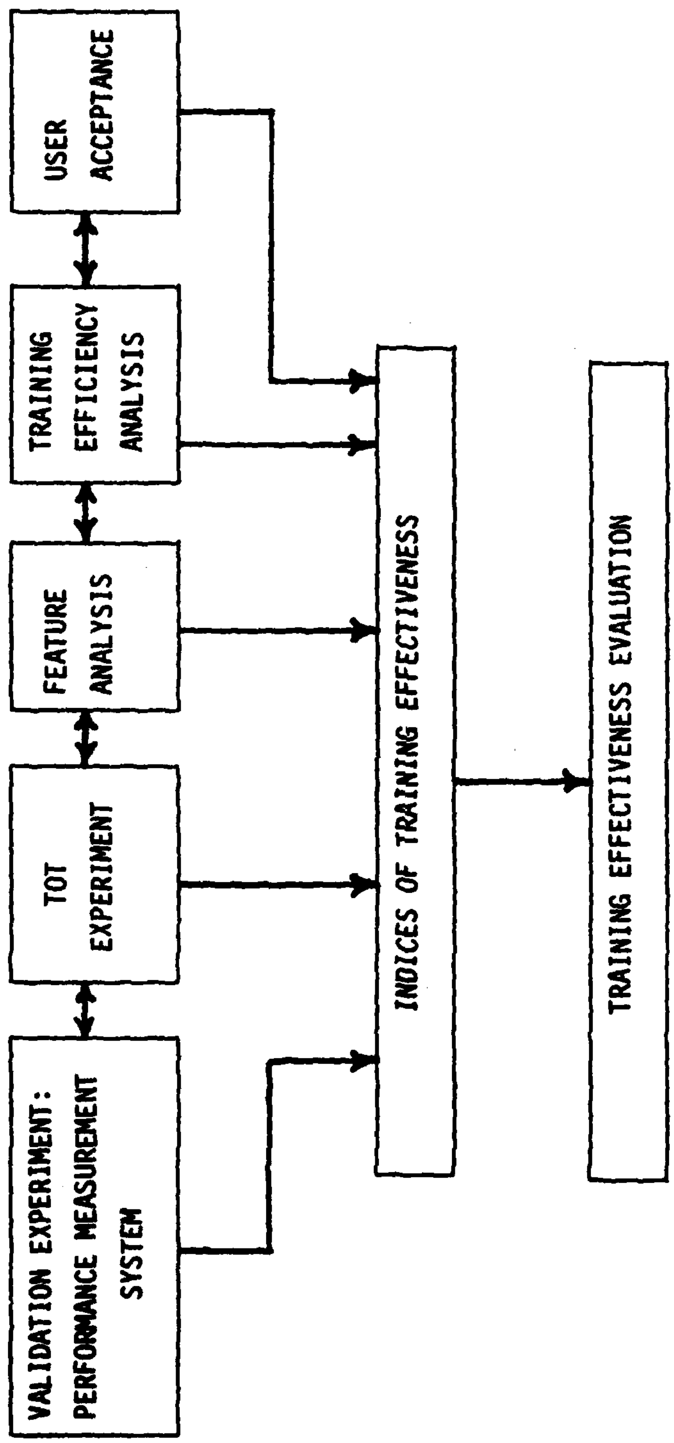


Figure 1. Organization of General Evaluation Approach

SECTION II

THE USE OF AUTOMATED SPEECH TECHNOLOGY IN TRAINING SYSTEMS

COMPUTER SPEECH RECOGNITION

Automated speech recognition systems are applicable to the training of tasks in which speech is the major component of task performance to be learned. Examples of these types of tasks within the Navy are Air Traffic Control (ATC), Landing Signal Officer (LSO) and Air Intercept Controller (AIC) tasks. All three involve a heavy emphasis on speech communication from the controller to a pilot. The speech is characterized by a set of standards, rules or procedures and a limited vocabulary. Current speech recognition technology can handle several hundred words or phrases (Dixon and Martin, 1979; Lea, 1980). This is an enormous reduction from the full range of the English language, but it provides adequate coverage for most operational tasks which use only a limited vocabulary.

Two important functions or outcomes of computer speech recognition are to control a simulation of the operational environment, and to serve as the input to a performance measurement system. An example of control through speech recognition is seen in the PARTS, where the controller gives "radio" transmissions to the pilot regarding azimuth and glideslope information. A simulated pilot/aircraft responds to this information and a corresponding change occurs on the simulated PAR display. In this way, the need for an additional human being in the system to support training can be eliminated, i.e. it is no longer necessary for a student to act as a simulated pilot.

Another function of computer speech recognition is the evaluation of trainee performance. Verbal behavior, as recognized by the computer speech recognition system, becomes the input for a performance measurement system. Student performance then can be evaluated, scored and diagnosed, and feedback can be provided to the student to give him information about his progress. This feedback function is an essential part of the learning process. Additionally, permanent, objective records of student performance may be stored as an output from the performance measurement system. All these functions are based on the initial input from the speech recognition system. Therefore, the accuracy of the speech recognition system is critical to the proper functioning of real-time control, performance measurement, feedback and record keeping, i.e., training control and management.

Current technology in speech recognition systems is primarily at the stage of isolated word recognition (IWR) or isolated phrase recognition. This type of system requires slight pauses between the phrases as well as sampling of the speaker's voice to establish reference patterns for recognition. This process is sometimes called "enrollment" or "voice training". The number of samples required may range from approximately 2 to 10 or more, for each phrase to be recognized. This task becomes a burdensome limitation of the technology when the vocabulary is large. For example, a student might have to provide speech samples up to 10 times for each of 200 phrases, or a total of 2,000 phrase samples in order to establish the data base for speech recognition. This can be time consuming and boring process which may not contribute directly to training after the first few repetitions.

Other potential limitations in the IWR technology include the limited vocabulary size, as mentioned previously, the recognition accuracy of the system, and finally, the speech stylization requirements. Stylization refers to the pauses between phrases which are required by the IWR technology. These types of systems are unable to parse continuous speech into its component phrases or words. Therefore, the student is required to make brief, artificial breaks between the phrases which are to be recognized by the computer. This type of stylization is not inherently difficult, but it does require the student to modify a highly overlearned behavior, i.e. normal speech. In critical control situations, for example, the student may tend to forget stylization, and speech recognition accuracy may suffer.

Since stylization is a learning process, it would be expected that long term use of the computer speech recognition system by a student would result in the greatest efficiency. Short term use of the system by an individual means that a considerable amount of time must be spent generating speech samples, and relatively little time may remain for actually using the speech recognition capabilities. Voice training time can be considered an investment, and longer term use of the system may yield greater return on that investment.

One potential benefit of computer speech recognition in training systems, however, may be derived directly from the voice training procedure. Tasks amenable to AST in training involve verbal output. In these types of tasks, verbal communications should be clearly articulated and highly practiced. Some of the requirements of a training system with AST, such as the speech sampling procedure and consistent stylized speech, are not incompatible with the job requirements of proper articulation and consistency. This certainly is the case in air traffic control. The issue becomes how to allocate students' time during training. Speech sampling may be beneficial for training, up to a point of diminishing returns. How to identify that point is not clear at the present time.

One of the primary benefits to be derived from the application of AST to training systems is the replacement of support personnel who normally would be required in the training setting. For example, the recipients of the students' verbal communications can be replaced by speech recognition and computer simulation. This can be a significant advantage in situations of manpower shortage, where additional personnel are required only to support the student's training, and they do not obtain significant training benefits in the support role.

Another primary benefit of the use of computer speech recognition in training is that it makes verbal tasks amenable to automated performance measurement. Previously, verbal tasks could be scored or evaluated only by an instructor. This is another area where AST offers the potential for enhancing training effectiveness and efficiency.

COMPUTER SPEECH SYNTHESIS

The technology of computer speech synthesis, or speech generation, is considerably more advanced than computer speech recognition because the technological requirements are less. Present synthesized speech is easily understandable. This capability allows the verbal output of another human being to be simulated in an automated system, complimenting the capabilities of computer speech recognition. In a training system for air traffic control, for example, computer speech synthesis can be used to simulate the voice of an instructor, pilot, or controller.

Computer speech synthesis may be regarded as another form of information presentation, an alternative to information presented on a CRT, hard copy printout or other types of visual or auditory information. Computer speech synthesis also could be used to provide warning or caution information and to simulate communications from another crew or team member with whom the student interacts. Additionally, there is great instructional potential for using speech synthesis to demonstrate proper phraseology, procedures, and techniques.

AST AND BEHAVIOR MODELING

The two components of automated speech technology complement each other well. Computer speech recognition simulates limited human recognition of verbal output from the student, and computer speech synthesis simulates the return of verbal communication from the person with whom the student was communicating. This verbal exchange requires an important additional step, however, which is modeling the behavior of the "other" person. For the case of air traffic controller, this other person is usually the pilot. Behavior modeling is necessary in order to connect speech recognition with speech synthesis in simulating the pilot.

SUMMARY

Computer speech recognition and synthesis provide new tools for training systems. However, automated speech technology per se, does not ensure good training systems. The use of such new technologies does not reduce the importance of good instruction and teaching techniques.

Computer speech recognition and synthesis provide the opportunity to develop training systems which reduce support personnel requirements and automate performance measurement for verbal tasks. Automated performance measurement of verbal tasks enables automation of record keeping and individualized instruction, thereby reducing instructor workload.

SECTION III

PARTS: BACKGROUND AND DEVELOPMENT

A brief description of the PAR task will be given, followed by a short discussion of the Navy's training program for Precision Approach Radar, and finally, a discussion of the development of PARTS and a description of the system.

PRECISION APPROACH RADAR

PAR and Air Surveillance Radar (ASR) are the two types of radar most commonly used in air traffic control. ASR provides bearing and range information (two dimensional) for broad area coverage, and has several uses besides air traffic control.

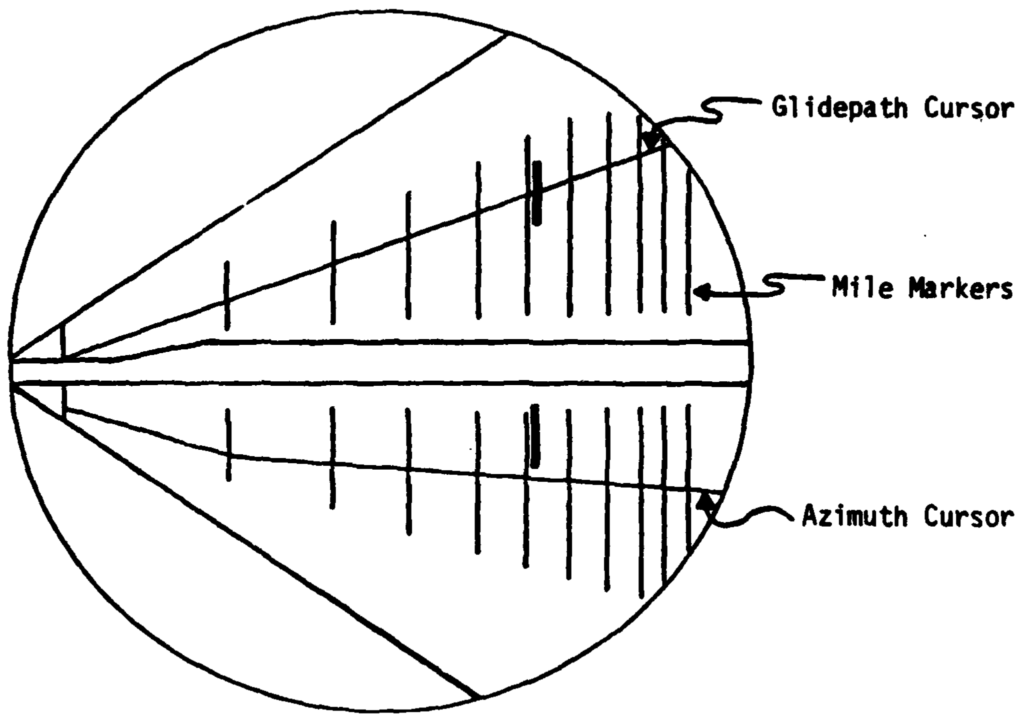
The term "precision" in PAR refers to the capability for providing a pilot with information about elevation (glidepath) as well as bearing and range during final approach to landing. PAR has limited area coverage, but can accurately determine an aircraft's position and movement in three-dimensional space. Therefore, it is well-suited for monitoring and guiding an accurate final approach to landing. Also, PAR does not require navagational aids (NAVAIDS) on the ground, or NAVAID equipment in the aircraft.

The PAR controller can serve as the "eyes" of the pilot during periods of low ceiling and visibility by giving him radio transmissions intended to maintain the final approach within safe limits. The control responsibility of the PAR controller ends when the pilot is informed that he is "at decision height", although the controller may continue to give advisories including "over landing threshold."

The PAR controller monitors the aircraft's position and trend on a radar scope that has two simultaneous display presentations, as shown in Figure 2. The upper portion represents elevation (glidepath) versus range, and the lower portion represents bearing (course) versus range. The example in Figure 2 shows an aircraft on glidepath, well right of course, slightly over 5 miles from touchdown.

The job of the PAR controller is important. In certain conditions, such as low visibility or aircraft instrument failure, his proficient job performance is critical to the safety of flight.

However, PAR is a subtask of the air traffic controller's job that has decreased in importance, except in the Navy. This trend away from PAR has been due to the use of instrument landing systems (ILS) and the development of more sophisticated automatic landing systems. Currently, the Navy uses PAR more than the other military services, and PAR is very rarely used in civilian aviation. The Navy practices PAR more frequently than other services and will continue to use PAR in the



Aircraft is on glidepath, well right of course, slightly over five miles from touchdown.

Figure 2. Typical PAR Display

foreseeable future because, in carrier operations, a form of PAR is used in the Carrier Controlled Approach.

NAVY PAR CONTROLLER TRAINING

The Navy Air Traffic Control (ATC) Schools, NATTC, conducts a 14 week training course "to provide selected aviation enlisted personnel with the basic control tower operator knowledge to meet the requirements of the Federal Aviation Administration for certification and the technical knowledge and skills which, when followed by practical experience, will lead to the fulfillment of the technical requirements for Air Controlman Third Class" (CNTECHTRA, 1976).

The tasks to be learned by the student in PAR training include equipment status checks and calibration, accepting control responsibility ("handoff") from the pattern controller, establishing radio contact with the aircraft, issuing corrective turn instructions to the pilot, transmitting glidepath and range advisories, advising the pilot when the aircraft reaches decision height, and conducting a waveoff or missed approach when necessary. In the conduct of these tasks the PAR controller monitors the radar scope, selects appropriate communications functions on an instrument panel, and transmits information to the pilot by radio.

The PAR portion of the curriculum currently is 5 1/2 days long. Classes of approximately 14 students receive one half day of classroom lecture on PAR procedures and phraseology, followed by four days of "hands on" practice using PAR laboratory simulators, concluding with a final performance test (P-run) on the sixth day. The PAR laboratory consists of 15 PAR consoles interfaced with a computing system. Aircraft responses are controlled by "bug operators" in another room. These are students who "fly" the simulated aircraft by pushing buttons on a Target Control Console (TCC) in response to the verbal transmissions by the PAR student. One bug operator is required for each student at a PAR console in the present simulation (Device 15G19).

The standard procedure is to assign half of a PAR class to be "bug operators" while the other half of the class is trained at the PAR laboratory consoles. The trainees switch roles for the next 50 minute session. There are approximaely 30 "bug operator" consoles in the Target Generation room, used for either ASR or PAR.

Students receive three weeks of training on ASR before entering the one week PAR training program. After PAR, the students receive three more weeks of training at the school.

PAR training at the ATC Schools is directed toward understanding and executing general PAR procedures. The student will not be qualified to conduct actual PAR final approaches until he has received extensive on-the-job-training (OJT) at a particular facility. Obtaining a PAR field qualification generally requires several months of OJT.

PARTS DEVELOPMENT AND DESCRIPTION

A conceptual feasibility study in 1973 recommended the development and construction of a demonstration model training system for the PAR phase of Ground Controlled Approach (GCA) Training (Feuge, Charles, and Miller, 1974). The PAR controller's task was deemed appropriate because a restricted set of verbal commands is used to achieve a well-defined goal. The PAR controller's verbal output is suited to the capabilities of automated speech recognition because of the limited and well-defined radio terminology promulgated by the Federal Aviation Administration (Breaux and Goldstein, 1975).

Logicon, Inc. began developing an automated PAR controller training system in 1974 under the sponsorship of the Naval Training Equipment Center (NAVTRAEQUIPCEN). The system was called Ground Controlled Approach-Controller Training System (GCA-CTS), and its development was documented in several technical reports (Barber, Hicklin, Meyn, Porter and Slemon, 1979; Breaux, 1976; Feuge, Charles, and Miller, 1974; Grady and Hicklin, 1976; and Hicklin, Nowell, and Petersen, 1978; and Hicklin, Barber, Bollenbacher, Grady, Harry, Meyn, and Slemon, 1980).

NAVTRAEQUIPCEN renamed the system Precision Approach Radar Training System in 1979 because the term GCA implied both ASR and PAR training. The system, however, was directed solely toward PAR. An experimental prototype PARTS was delivered to NATTC, NAS Memphis, in the Fall of 1979.

The following major functions are automated in PARTS: speech recognition; speech generation; simulation of aircraft and pilot performance; computer assisted instruction; model controller demonstrations; syllabus control; and student performance measurement and record keeping. The experimental prototype PARTS has only one student station and one instructor station, but the possibility for multiple student stations is basic to the system design. PARTS is divided physically and functionally into three areas: student station; instructor station; and system control.

The student station has a CRT (Data General 6053) for presenting alphanumeric instructional material; a keyboard for non-verbal interaction with the system; a communications panel for radio, ICS, and clearance request; and a computer-generated PAR display (Megatek MG552) of elevation, azimuth and range, similar to the example shown in Figure 2.

The instructor station, located in another room, has three major components: a CRT (Data General 6053) with keyboard; audio communications panel; and a serial character printer (Tally 1602).

The third functional area, system control, consists of the following equipments mounted in a twin-bay equipment rack: two minicomputers (Data General, Eclipse S/130); a 10 megabyte cartridge disk; dual floppy discs; speech synthesizer (Federal Screw Works, VOTRAX VS 6.4); and voice input preprocessor (Threshold Technology, Threshold 500).

An overview of the training system is given by Logicon (1979), from which Figure 3 was obtained. It depicts the relationship of the following functional features of the system: CAI; voice generation; speech recognition; simulated PAR display; aircraft/pilot/environmental (APE) model; PAR model controller; performance measurement and evaluation; performance feedback; record keeping; and syllabus control.

When the PARTS syllabus was being developed initially in 1974-1976, the ATC Schools syllabus alloted 10 days for PAR training. Subsequently, this was reduced to 5 days, causing corresponding reductions in the PARTS syllabus. The experimental prototype PARTS has been designed to provide basic training in PAR procedures with the goal that competent students can complete the course within the five and one half days currently allotted to PAR training.

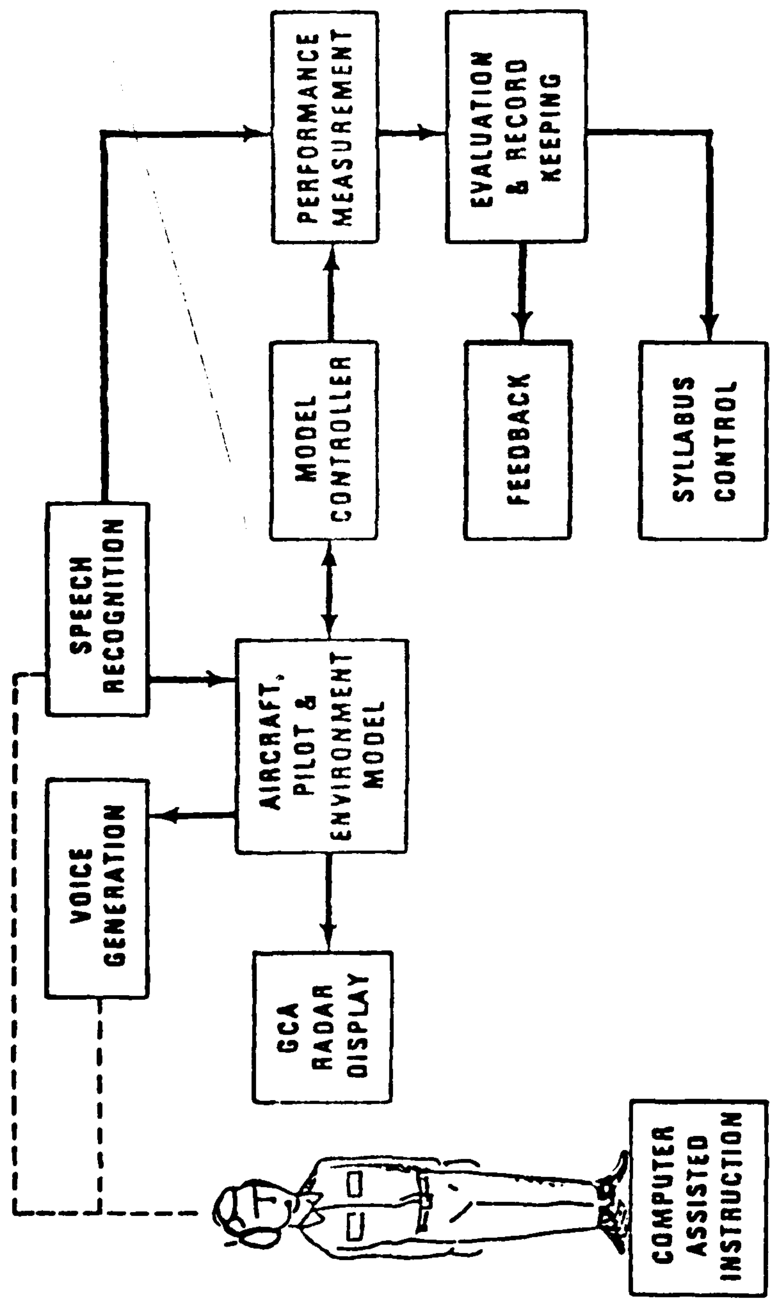


Figure 3. PARTS Training System Diagram (from Logicon, 1979)

SECTION IV

OVERVIEW OF DATA COLLECTION METHODS AND PROCEDURES

The methods used for data collection in the PARTS evaluation included review of literature and documentation, observation, questionnaire survey, interview, and experimentation.

LITERATURE AND DOCUMENT REVIEW

Three main categories of literature and documentation were reviewed and analyzed: training effectiveness evaluations (TEE); development of GCA-CTS (PARTS); and NATTC Air Traffic Control Schools curriculum and associated course materials. The TEE literature was reviewed for techniques and procedures, particularly transfer of training designs, that could be applied effectively to the PARTS evaluation. All NAVTRAEQUIPCEN and Logicon published reports which pertained to the research and development of GCA-CTS were examined. The logic underlying design decisions was noted, as well as the reasons for changes in the system during its development cycle. Documents were obtained from the school pertaining to the curriculum, programmed texts, terminal and enabling objectives for the PAR and ASR courses, grading criteria and sample grading sheets for the PAR laboratory final examination.

OBSERVATION

A two-day preliminary observation period at Logicon in San Diego enabled members of the Canyon project team to obtain initial familiarity with PARTS. Two students from ATC School used the system in this preliminary trial period.

The first major period of observation was 6-21 November, 1979 when the system was observed at the ATC Schools. Detailed notes were taken of the interaction of students and instructors with the system. Instances of student confusion or uncertainty were noted. The frequency of selection of instructional options was recorded as well as the assignment of system-determined variables. During the on-site period, observation was conducted nearly full-time by either the Canyon Research Analyst or Principal Investigator. The observer was seated behind the student at the student's console, providing an excellent opportunity for observing the student's interaction with the system.

The second major observation period was during the transfer of training study, from 10 March through 23 April, 1980. The same techniques were used for data collection with the addition of hand-recording a sample of student transmissions and system responses on over 70 practice problems.

INTERVIEW

Interviews were conducted with students who had completed the PARTS performance test (P-run), and with the two instructors assigned to PARTS. A substantial amount of the information in the present report was obtained from interviews with the PARTS instructors because their experience spans standard PAR training, PARTS, and the actual operational environment of the control tower.

QUESTIONNAIRE SURVEY

Questionnaires were developed to survey students' expectations about PARTS, (Expectations Questionnaire) and their general background and attitudes toward the military and ATC School training (PARTS Student Questionnaire). PARTS students who participated in the Transfer of Training Study gave feedback about their training in the Post Training Questionnaire.

PERSONNEL RECORDS

Existing measures of student aptitude and achievement were obtained from the students' permanent records or ATC School records. These measures included the Armed Services Vocational Aptitude Battery (ASVAB), ATC School grade average, Reading Comprehension Test, and Otis-Lennon Mental Abilities Test. All such data were treated as confidential.

PERFORMANCE MEASUREMENT VALIDATION STUDY

The PARTS performance measurement validation study was designed to test the automated performance measurement system's ability to discriminate between experienced PAR controllers and students who were naive with respect to PAR. Initially, the procedures called for obtaining PARTS P-run scores for small groups of students and instructors during a three week period. Subsequently, it became clear that the P-run could not be administered without an extended period of time (about 30 hours) required for each person to progress through the PARTS syllabus. It was not feasible for either instructors or students to contribute 30 hours of their time without severely interfering with normal training at the ATC Schools.

A short syllabus was developed to enable participants in this study to progress rapidly to the P-run. Approximately 6 hours was required for minimum voice data collection, brief instruction on procedures, and a small number of practice trials. Three P-runs rather than one were included in the short syllabus in order to conform with the standard PAR lab P-run.

TRANSFER OF TRAINING STUDY

One criterion of training effectiveness is the performance level of the students after training. A transfer of training (TOT) study was designed to compare the performance of two groups of students, PARTS-trained and traditional PAR, on a criterion test. The objective was to compare TOT for the two groups of differentially trained students for the purpose of providing an index of the relative effectiveness of the training methods. The criterion test included a set of three simulated PAR approaches designed to be as similar to an actual field PAR task as possible. A stratified sample of students (High, Medium, and Low) was assigned to the two training methods based on ATC School grade average prior to PAR. This measure was found to be the best predictor of PAR grades in a pilot study. The criterion test was administered on the standard PAR laboratory equipment, and a brief transition training period was given to both groups of students before testing.

GENERAL PROCEDURE FOR PARTS USE

The general procedure for the use of PARTS during the evaluation period was to assign two students to the system each week. The two PARTS instructors each took responsibility for one of the students. The instructors and students worked split six-hour shifts on the system each day. This schedule was not in effect from mid November through mid January, when the Christmas break and the Performance Measurement Validation (Short Syllabus) Study caused a change in the routine.

A total of 34 students participated in PARTS training to some degree. The course was completed, including the P-run, by 24 students. Canyon representatives directly observed portions of the training of 16 students.

SECTION V

AUTOMATED SPEECH TECHNOLOGY IN PARTS

CONTROLLING A PAR APPROACH

In PARTS, one of the primary uses of computer speech recognition (CSR) was to enable the student to control a simulated PAR approach without an actual pilot, aircraft, or the requirement for another student to simulate the pilot. This was accomplished by the subsystem Aircraft/Pilot/Environmental Model (APE). The student controller's verbal transmissions, such as a call for a turn, were recognized and then acted on by the APE model, resulting in a turn of the aircraft on the PAR display. This is, in effect, a closed loop control system from the student to the APE model back to the student via the display. The student controls two major variables in PAR, as discussed previously, by giving turn instructions (vectoring), and glidepath position and trend messages. Additionally, waveoffs or missed approaches can be called for. Many other messages or phrases are given by the PAR controller, but most of these are advisory in nature and not directly controlling in the sense of causing a response on the part of the pilot.

The use of CSR for approach control in the PARTS represents a major advance relative to current and traditional training methods. Currently, the pilot is simulated by another student operating a push button console. This student position is the target control console operator, better known as "bug operator", who listens to the student controller's verbal transmissions and "flys" the aircraft accordingly. A student in PAR spends roughly half of his training time flying the bug. This represents unproductive training time. Very little training value is gained by bug operation. Consequently, the use of CSR and APE effectively doubles productive training time.

A related use of CSR is to simulate a pattern controller. The pattern controller gives the initial handoff to the PAR controller and, after a touch-and-go or a missed approach, the PAR controller gives control of the aircraft back to the pattern controller. These handoff procedures are not frequently practiced in the current training at the ATC Schools. Occasionally the instructor will act as pattern controller to give the student practice in accepting handoffs, but generally the PAR student controls approaches from about 7 to 9 miles from touchdown without a handoff. The approach is started simply by telling the bug operator to begin. Therefore, PARTS provides the capability for more realistic training in terms of the handoff procedures and interaction between the PAR controller and the pattern controller.

SPEECH RECOGNITION AND PERFORMANCE MEASUREMENT

Another function of computer speech recognition in PARTS was to provide input to the automated performance measurement system. Speech is the primary output of the PAR controller, and therefore, his task previously has not been amenable to automated performance measurement. An instructor always has been required to evaluate student verbal performance. Computer speech recognition provides input to the performance measurement system, which relieves the instructor's burden of scoring and evaluating trainee performance on each approach. The performance measurement subsystem, in turn, provides information for performance feedback, record keeping, and syllabus control.

VOICE TRAINING

Table 1 lists the 107 phrases included in the PARTS syllabus.

The term Voice Training was used in PARTS to refer to the enrollment procedure, or speech samples, given by the student to build up a reference library within the computer. Some phrases were repeated 4 times and some 10 times before they could be used. The total time required for the voice training procedure was difficult to estimate, since Logicon nicely integrated the voice training procedure with instruction and learning of PAR procedures. Voice training was spread throughout the seven levels of the syllabus (see Section VIII of this report).

Speech recognition is more accurate when voice training is done in the context in which the phrases will be spoken. This was a problem in PARTS because voice training frequently was done with the student facing the small CRT where a phrase was displayed as a prompt for the student to speak it. Subsequent prompts included a replay of the student's digitized voice, or use of the computer speech synthesis system. However, voice training often was done outside of the context of a PAR approach. Speech recognition, on the other hand, was attempted with the student attending to the simulated PAR display and controlling an approach. The context effect of voice training probably is due to inflection and emphasis induced by the intended effect of the phrase. Repeatedly saying the phrase in response to prompts is different from saying a phrase in the context of controlling an aircraft.

There may be some room for improvement in PARTS by carefully examining the possibilities of contextual speech sampling. This is not an easy problem to solve because the student simultaneously is attempting to learn procedures for use of the phrase. Possibly the phrase could be taught initially through demonstrations, then updated speech samples could be obtained while the student is observing part of the approach appropriate to the use of the phrase. The students were instructed to attempt to say the phrases as though they were controlling the aircraft, but this requires acting ability as well as prior knowledge of the use context.

TABLE 1. PARTS SPEECH RECOGNITION PHRASES

Phrase Number	Phrase
1)	1 MILE
2)	1 AND 1/2 MILES
3)	2 MILES
4)	2 AND 1/2 MILES
5)	3 MILES
6)	3 AND 1/2 MILES
7)	AT
8)	TWELVE
9)	FIFTEEN
10)	TWENTY
11)	TWENTY-FIVE
12)	THIRTY
13)	0
14)	1
15)	2
16)	3
17)	4
18)	5
19)	6
20)	7
21)	8
22)	9
23)	CONTACT TOWER AFTER LANDING
24)	BUTTON ONE CLEAR
25)	BUTTON TWO CLEAR
26)	MISSED APPROACH
27)	IF RUNWAY NOT IN SIGHT
28)	IF RUNWAY NOT IN SIGHT EXECUTE MISSED APPROACH
29)	IF RUNWAY NOT IN SIGHT CLIMB AND MAINTAIN ONE THOUSAND FIVE HUNDRED
30)	BUTTON ONE
31)	PROCEED DIRECT POINT BRAVO HOLD UNTIL ADVISED BY GCA
32)	BUTTON TWO
33)	ON THE GO
34)	OVER LANDING THRESHOLD
35)	TOO FAR LEFT FOR SAFE APPROCH
36)	TOO FAR RIGHT FOR SAFE APPROACH
37)	ON CENTERLINE

TABLE 1. PARTS SPEECH RECOGNITION PHRASES (CONTINUED)

Phrase Number	Phrase
38)	LEFT OF CENTERLINE
39)	SLIGHTLY LEFT OF CENTERLINE
40)	RIGHT OF CENTERLINE
41)	SLIGHTLY RIGHT OF CENTERLINE
42)	TOO LOW FOR SAFE APPROACH
43)	TOO HIGH FOR SAFE APPROACH
44)	WIND
45)	CLEARED FOR LOW APPROACH
46)	CLEARED FOR TOUCH AND GO
47)	CLEARED TO LAND
48)	1 MILE FROM TOUCHDOWN
49)	2 MILES FROM TOUCHDOWN
50)	3 MILES FROM TOUCHDOWN
51)	4 MILES FROM TOUCHDOWN
52)	WELL LEFT OF COURSE
53)	LEFT OF COURSE
54)	WELL RIGHT OF COURSE
55)	RIGHT OF COURSE
56)	WELL BELOW GLIDEPATH
57)	WELL ABOVE GLIDEPATH
58)	GOING FURTHER BELOW GLIDEPATH
59)	GOING FURTHER ABOVE GLIDEPATH
60)	CLIMB AND MAINTAIN ONE THOUSAND FIVE HUNDRED
61)	AT DECISION HEIGHT
62)	ON COURSE
63)	SLIGHTLY LEFT OF COURSE
64)	SLIGHTLY RIGHT OF COURSE
65)	CORRECTING
66)	ON GLIDEPATH
67)	BELOW GLIDEPATH
68)	SLIGHTLY BELOW GLIDEPATH
69)	ABOVE GLIDEPATH
70)	SLIGHTLY ABOVE GLIDEPATH
71)	GOING BELOW GLIDEPATH
72)	COMING UP
73)	GOING ABOVE GLIDEPATH
74)	COMING DOWN
75)	POSITION 4 ROGER
76)	RADAR BUTTON 1
77)	RADAR BUTTON 2

TABLE 1. PARTS SPEECH RECOGNITION PHRASES
(CONTINUED)

Phrase Number	Phrase
78)	THIS IS YOUR FINAL CONTROLLER HOW DO YOU HEAR ME?
79)	WHEELS SHOULD BE DOWN
80)	DO NOT ACKNOWLEDGE FURTHER TRANSMISSIONS
81)	APPROACHING GLIDEPATH
82)	BEGIN DESCENT
83)	GIVE ME BUTTON ONE
84)	GIVE ME BUTTON TWO
85)	ARMY EIGHT SEVEN SIX
86)	MARINE SIX EIGHT SEVEN
87)	NAVY THREE ONE ZERO
88)	AIR FORCE THREE ZERO SEVEN
89)	OVER
90)	THIS WILL BE A NO-GYRO APPROACH
91)	MAKE HALF STANDARD RATE TURNS
92)	5 MILES FROM TOUCHDOWN
93)	6 MILES FROM TOUCHDOWN
94)	7 MILES FROM TOUCHDOWN
95)	8 MILES FROM TOUCHDOWN
96)	LOW ALTITUDE ALERT CHECK YOUR ALTITUDE IMMEDIATELY
97)	HOW DO YOU HEAR ME NOW?
98)	CORRECTION
99)	TURN RIGHT
100)	STOP TURN
101)	TURN LEFT
102)	EXECUTE MISSED APPROACH
103)	RADAR CONTACT LOST
104)	CLIMB AND MAINTAIN THREE THOUSAND
105)	TURN RIGHT HEADING
106)	HEADING
107)	TURN LEFT HEADING

The point has been made that, due to this initial transition period of voice training, the benefits of AST training would be more apparent in a course longer than the 5-day PAR course. However, one individual pointed out a possible benefit of the "learning to speak to the machine" training time. Eventually, voice technology may occur in the applied setting, and students are much more likely to feel comfortable with (and to use) an innovative technology in an operational environment if they have been introduced to it in training. Experienced personnel who have been trained by traditional methods and have operated for many years without voice technology, may be more reluctant to accommodate these new technologies.

RECOGNITION ACCURACY VARIABLES

Recognition accuracy is critically important for speech recognition systems in training. However, recognition accuracy cannot adequately be described simply by total percent accuracy. The importance of recognition accuracy dependents upon categories of errors. Certain types of errors are critical, causing disruption of the training situation and loss of training time. Other types of recognition errors are inconsequential. Before presenting data on recognition accuracy in PARTS, a brief discussion of the variables influencing recognition accuracy will be given.

A partial list of student variables that may influence computer speech recognition accuracy is as follows: microphone level control; speech stylization; speech consistency; student learning; stress; and fatigue. PARTS students were instructed to observe the speech volume meter. The needle should have pointed to 5 (mid-range) during their speech. The students were instructed to manipulate the control volume to achieve this mid-range level during voice training and each time they sat down to the training system thereafter. Despite these reminders, given in the Student Guide and presented on the CRT, some students were observed not heeding these precautions. Some spoke very loudly at the system, apparently feeling that they had to shout at it in order to be understood. This practice undoubtly decreasesd the recognition accuracy of the system.

Consistency is another variable that is important in recognition accuracy. Consistency may refer to loudness, inflection, or sytlization/pauses. The stylization requirements are not particularly difficult but, as mentioned previously, they conflict with the highly over learned practice of normal speech. Therefore, students often were inconsistent in speech stylization during the first day or two on the system. Recognition accuracy tended to increase about the third day on the system, even though the vocabulary size increased each day. Undoubtedly, the reason for this effect is learning. It simply takes time on the system, apparently 4 to 12 hours or so, depending on the student, in order to achieve consistency in speech level, inflection and stylization. As in all skills, individuals differ in ability. Some students quickly achieved good recognition accuracy and consistency in speech, while other students never achieved excellent speech recognition accuracy because of their failure to learn speech stylization and consistency. Enhanced effectiveness of speech recognition systems in the future will have the largest benefit for students who are weaker in learning the speech stylization requirements. Advances in the technology should allow greater flexibility of input, placing the burden on the system rather than on the student to conform to speech stylization rules.

Two other variables that seemed to influence recognition accuracy were stress and fatigue. These variables are difficult to measure and in the PARTS evaluation we did not attempt specifically to measure them. The instructors reported, however, that speech recognition often tended to decrease in accuracy toward the end of a six hour session, as the student became fatigued. Personal communication from (Connolly, 1979) indicated that the fatigue-induced decrease in recognition accuracy has been observed in his work with the FAA. In fact, he has used two separate voice training data banks for recognition of his own voice. One data bank or voice sample taken when he was fresh, and one when he was fatigued. Similar procedures for updating speech samples have been used when speech recognition deteriorated due to colds or allergies (Connolly, 1979).

Stress also has an effect on recognition accuracy, although the entire PAR training was not a particularly stress inducing situation. No student failed, and students seemed to understand that anyone participating in this project was not going to fail. Transient stress seemed to be induced in certain approach control situations where the student was having difficulty, or where a speech recognition error lead to an unusual aircraft maneuver, setting up a very difficult approach control situation. Students frequently would forget to pause between transmissions in these cases. In their rush to regain control of the approach they would forget stylization and change the volume and inflection of their voices, leading to decreased speech recognition accuracy. Observation of highly experienced controllers, such as the instructors, showed that they had the discipline and presence of mind, developed through years of experience, to regain control of the approach while maintaining speech stylization constraints. Remaining calm under stress is an important characteristic of a skilled air traffic controller, but it is doubtful that it can be taught in an introductory, five day PAR course.

Two variables that appear not to influence speech recognition were sex and regional accent. Males and females had about the same recognition accuracy, and regional accent seemed to make no difference in recognition accuracy. One student, who's first language is Spanish, spoke with a distinct accent but this caused no speech recognition problem. The key concept appears to be consistency. As long as the student pronounces the phrase consistently, speech recognition is not affected by variables such as sex (voice pitch) or accent.

RECOGNITION ACCURACY DATA

Three types of data were collected on recognition accuracy. Percent recognition accuracy was derived from computer printouts of the students' performance tests (P-runs), the frequency of voice retraining for the 107 PARTS phrases was complied, and system responses to student transmissions were observed in 71 practice problems.

P-RUN TRANSMISSION RECOGNITION ACCURACY. Computer print-outs of performance data for 22 of the students who completed the P-run were examined for student transmissions that were unrecognized or misrecognized by the speech understanding system. Transmissions rather than phrases were used as the unit of measurement because they are the functional unit for the PAR task. Many transmissions did consist of one PARTS phrase, such as "On Glidepath" or "Four miles from touchdown", but some consisted of several PARTS phrases, such as "Marine 687, turn right heading 165, over", which has six phrases. Each digit in a vector is a phrase.

A computer record of each P-run was used by the instructor during the Modify procedure. The instructor listened to a replay and wrote the actual transmission adjacent to the automated recognition whenever any discrepancy was found. Table 2 presents the transmission recognition accuracy (TRA) data for 22 students obtained from the Modify procedure.

The overall TRA for 1546 P-run transmissions was eighty-five percent. The range of TRA was from fifty to ninety-seven percent. However, the fifty percent TRA occurred immediately after the elimination of software that was designed to enhance speech recognition. Canyon interviewed the student the following day, and he commented that he had no problem with speech recognition until the P-run. Subsequent students, number 10-22, had better recognition accuracy with the modified software, as can be seen in Table 2. Perhaps student number 10, with fifty TRA, was hampered by training for four days with one type of speech understanding software, then given the P-run with another type.

Many of the unrecognized transmissions occurred when students failed to pause between phrases. Some of these cases were obvious simply by inspecting the P-run record. For example, a student getting nearly perfect recognition would have two consequtive unrecognized phrases like "Going further below glidepath, well below glidepath". Some students had a tendency to omit the pause between phrases like these. In this example, two transmission errors would be scored because the failure to use proper stylization created one lengthy phrase that the IWR system could not recognize.

The major problem evident in the P-run transmission recognition errors was the relatively common occurrence of a vector misrecognition of large angular magnitude. For example, within a few degress of 160^{o} were recognized as turns to 300^{o} a total of 14 times, during 11 of the 22 P-runs. On at least five P-runs, control of the approach was lost as a direct consequence of this misrecognition. This problem seemed to occur only when the student would give a vector at a range of approximately one mile from touchdown. Perhaps, a simple software bug was responsible for this problem, but this type of disruption of approach control is very undesireable and must be avoided in an operational system.

TABLE 2. TRANSMISSION RECOGNITION ACCURACY (TRA) ON P-RUNS

STUDENT	DATE	NUMBER OF ERRORS		NUMBER OF TRANSMISSIONS		TRA (% CURRENT)	FALSE 300° VECTORS
1	10/24/79	13		117		89	
2	11/6	14		86		84	
3	1/22/80	14		104		89	
4	2/4	6		58		90	
5	2/4	4		54		93	
6	2/19	7		54		87	
7	2/26	6		67		91	
8	2/26	9		74		88	1
9	3/11	13		71		82	
10	3/11	27		54		50	
11	3/11	8		68		88	1
12	3/18	18		57		68	1
13	3/25	18		85		79	1
14	3/25	14		53		74	
15	4/1	8		58		86	3
16	4/1	2		72		97	
17	4/8	7		64		91	1
18	4/8	6		67		91	1
19	4/15	6		64		91	1
20	4/15	11		76		86	1
21	4/22	12		75		85	1
22	4/22	11		68		84	2
	TOTAL	234	/	1546	=	85%	14

VOICE RETRAINING FREQUENCY. Voice retraining refers to the selection of Voice Test by the student to test recognition accuracy on a particular phrase, and the selection of the INIT NEW R/T function, which allows the student to select a particular phrase or set of phrases for voice retraining. Voice retraining is a partial replacement of the criterion set of reference phrases for a particular student.

Frequently, improvements in recognition accuracy of a particular phrase were observed after the use of the INIT NEW R/T function (voice retraining). This effect would be expected if the student was not speaking naturally the first time he or she trained the phrase or, if he changed the way he said the phrase over time. An example of this was a student who trained the phrase "one and one half miles from touchdown" and later used the phrase "one and a half miles from touchdown".

The availability of Voice Test was limited in the PARTS design, and access to voice retraining was not available to the student at the beginning of the PARTS evaluation. Greater access to these functions was given to the student through a software change requested by the PARTS instructors. This appeared to be a very worthwhile change. Students were knowledgeable and responsible to use these functions for improving phrase recognition accuracy.

From automated records, Canyon compiled the retraining frequency of the 107 phrases used in PARTS by 17 students. These data indicate which phrases either were most difficult for the system to recognize or, were most difficult for the students to repeat consistently. Twenty six of the 107 phrases were never selected for retraining. Conversely, seventy-six percent of the phrases were selected for retraining at least once. The total number of times a phrase was retrained ranged from 1 (on 20 phrases) to 15 (on two phrases, "three miles from touchdown," and "turn right heading"). Range calls were selected for retraining frequently. "One mile from touchdown" had 8 retraining selections, "two miles from touchdown" had 14, "three miles from touchdown" had 15, "four miles from touchdown" had 14, and "five miles from touchdown" had 13. The difficulty in recognizing range calls probably stems from the fact that these relevately long phrases "___ miles from touchdown" differed only in the first word.

Other phrases that were selected frequently for voice retraining included "turn right heading" and " turn left heading", with 10 selections for retraining each. Glidepath position also was retrained frequently . Specifically, "slightly below glidepath" with 10, "above glidepath" with 5, "slightly above glidepath" 10, and "going below glidepath" 7. Other phrases causing some difficulity with recognition were digits: "zero" was retrained 6 times, "one" was retrained 8 times, "five" was retrained 6 times, "six" was retrained 10 times, and "eight" was retrained 7 times. One reason that some digits seemed difficult to discriminate is that they are so short. But, certain digits were rarely selected for retraining. Examples are the digit "two" which was retrained once, "three" was never retrained, "seven" was retrained only once, and "nine" was retrained only once.

Several factors may have influenced the selection of voice retraining, such as frequency of phrase use and phrase importance. The frequency of use of the phrase was a factor because several phrases such as, "this is your final controller how do you hear me", were used on every approach. Certain phrases were used less frequently and, therefore, the speech recognition error frequency would have been lower. Similarly, the importance of the phrase was relevant to voice retraining. Students could not maintain approach control without using certain phrases such as "turn right heading" and "turn left heading". Therefore it was essential that these phrases be retrained if there were any problems in speech recognition accuracy.

SPEECH RECOGNITION IN OBSERVED APPROACHES. Canyon personnel observed many PARTS practice approaches during the course of this evaluation. On 71 approaches, the observer logged the students' transmissions and the response of the system. The controller transmissions that resulted in an observable response from the system included: 1) vectors (turn commands) resulting in a change in direction of the aircraft as displayed on the simulated PAR scope; 2) most transmissions given during the first four miles of the approach. Prior to the transmission "do not acknowledge further transmissions," which is given at approximately 5 miles from touchdown, controller transmissions received a simulated verbal response from the pilot. These verbal responses provided the observer with information about speech recognition errors. It should be noted, however, that failure of the system to give a verbal response could be due to system errors other than the speech recognition system, or due to the student failing to unkey the microphone after transmitting. Thus, the lack of a synthesized verbal response was not necessarily a direct indicator of speech recognition failure.

Vector errors were easily observable prior to "do not acknowledge..." because the system responded both by turning the aircraft and repeating the transmission. An example from level 6 of the syllabus on 1 April 1980: Student - "Marine 687, turn left heading 160", System - "Roger, turn left heading 145". And, simultaneously, the aircraft displayed on the scope began a turn to 145°.

Some speech recognition errors were impossible to detect by observation, such as when the transmission did not require a response of the aircraft, and the simulated pilot merely replied "Roger". Course information is an example of this category: Student - "Well right of course, correcting", System - "Roger". This provides no information about whether the transmission was recognized correctly. Therefore, these kinds of transmissions were excluded from the data base.

Table 3 provides a sample of the speech recognition data obtained form each approach. The starred items were judged as errors. No distinction was made in the data between unrecognized phrases, when the system gave no response, and misrecognized phrases, when the system responded incorrectly. In the example, the two errors out of a total of nine transmissions yielded a transmission recognition accuracy (TRA) score of seventy-eight percent.

TABLE 3. SPEECH RECOGNITION IN AN OBSERVED PRACTICE PROBLEM

Transmission Number	Student Controller Transmission		System Response
1	Army 876, Radar Button 2		Released button
2	Army 876, This is your final controller, how do you hear me?		Loud and clear
3	Army 876, Turn left heading 170		Roger, Turn left heading 170
4	Army 876, Wheels should be down		Roger, Wheels down and locked
5	Army 876, Turn left heading 160		Roger, Turn left heading 160
6	Army Approaching glidepath	*	Roger, Turn left heading 160
7	Army 876, Approaching glidepath	*	Roger, Turn left heading 160
8	Turn left heading 158		Made turn
9	Turn left heading 154		Made turn

*Judged as errors in system response

The 71 problems observed were compiled both by student and by syllabus level. Table 4 lists the TRA data for five students, ranging from sixty-nine to eighty-six percent. Table 5 lists the TRA by syllabus level. The effects of student and syllabus level are confounded but the data give some indications of the variability among students. For example, two students were observed in syllabus level 2. Student B had a TRA of eighty-one percent while student E had a TRA of fifty-one percent in syllabus level 2.

TABLE 4. TRANSMISSION RECOGNITION ACCURACY (TRA) BY STUDENT

Student	Number of Approaches	Error Freq. by Type 1	2	3	4	Errors/Total	TRA(%)
A	9	8	7	1	3	19/81	76
B	5	4	5	3	2	14/45	69
C	44	33	50	8	14	105/426	75
D	9	7	7	0	1	15/109	86
E	4	6	3	0	1	10/17	41
					Total	163/678	76%

TABLE 5. TRANSMISSION RECOGNITION ACCURACY BY SYLLABUS LEVEL

Syllabus Level	Number of Approaches	Error Frequency by Type 1	2	3	4	Errors/Total	TRA(%)
2	10	8	9	1	3	21/76	76
3	10	11	11	1	2	25/104	80
4	15	14	21	3	5	43/147	71
5	0						
6	25	17	24	6	7	54/255	79
7	8	5	7	1	2	15/65	77
P-run	3	3	0	0	2	5/40	88
					Total	163/678	76%

The 163 observed recognition errors were classified into four categories: 1) no problem; 2) minor problem in controlling the approach; 3) major problem in approach control and 4) control of approach was lost. Two examples of each category of recognition error are shown in Table 6.

TABLE 6. EXAMPLES OF FOUR CATEGORIES OF RECOGNITION ERRORS

	Error Type	Student Transmission	System Response
No Problem	1	Marine 687, Radar Button 1	None
	1	Air Force 307, wheels should be down	None
Minor Prroblem	2	(Aircraft heading 162^{o}) Marine 687, turn right heading 164	None
	2	(Aircraft heading 160^{o}) Army 876, Approaching Glidepath	Roger, turn left heading 160
Major problem	3	Air Force 307, turn left heading 160	Roger, turn left heading 145
	3	Navy 310, turn left heading 145	No response (second consecutive error in increasingly bad situation)
Approach Control Lost	4	"Give me button 1" (6th consecutive identical call)	No response (student did not acquire control of approach. System waved-off)
	4	(Aircraft heading 155^{o}) Air Force 307, turn right heading 160^{o}	Roger, turn right heading 300

The examples are given in the table to aid in defining the four categories of recognition errors. The frequency of the four types of errors was shown previously in Tables 4 and 5 by student and by syllabus level, and is summarized in Table 7.

TABLE 7. FREQUENCY OF RECOGNITION ERROR TYPES

	Error Type				
	1 No Problem	2 Minor Problem	3 Major Problem	4 Lost Control	Total
Frequency	58	72	12	21	163
Percentage	36%	44%	7%	13%	100

The results shown in Table 7 indicate that eighty percent of the transmission recognition errors caused only minor problems or no problems during the observed practice approaches. On the other hand, control of the approach was lost on thirteen percent of the recognition errors. Another way to assess the impact of recognition errors on approach control is that control was lost in 21 of 71 approaches sampled, or thirty percent. This situation usually occurred when the pilot/aircraft would initiate a radical turn (presumably to 300°) and quickly depart from the final approach course. When that occurred, the approach eventually was terminated either by the system, the student or the instructor.

A single index of speech recognition accuracy, such as percent correct, is insufficient for evaluating CSR in training because of the varied importance of speech recognition errors. Certain errors were trivial, such as a failure to recognize a "slightly above glidepath" transmission. This had no meaningful effect on the approach, and was a phrase that simply would be repeated by the student controller. Other phrases, particularly vectoring ("turn right heading 160") were essential for controlling the approach. A vector misrecognition by PARTS was not too serious if the recognized heading differed from the spoken one by only a few degrees. But when the recognized heading was greater than approximately 15° in error, the student had a control problem unless a countercorrective turn was given promptly, and correctly recognized.

A speech misrecognition involving a large magnitude vector error perturbs the approach considerably. For example, it was not uncommon to observe the system recognizing a vector in the 150° to 170° range as "turn right heading 300" (the runway heading in PARTS was 160°). This represents an angular error of at least 130°. Often approach control was lost when such a misrecognition occurred. The aircraft would make a hard right turn, and within a matter of seconds, be off the simulated PAR display. Students would attempt to regain control by giving increasingly radical counter vectors. Sometimes, control was successfully regained, but often such extraordinary attempts were not understood, probably because the student was changing speech patterns. This situation represents the most severe criticism of computer speech recognition in PARTS. Such erratic pilot/aircraft behavior is unrealstic because an actual pilot would never take such a vector. Training is disrupted because the approach usually is terminated, preventing practice on subsequent procedures, such as landing threshold sequence. Also, this situation violates the concept of incremental training because the student is introduced to a very difficult problem beyond his control capability.

Fortunately, this type of problem with computer speech recognition in PARTS should not be too difficult to overcome. Software enhancements to the speech recognition system could be improved for rejecting recognition of vectors with excessive angular error. This type of strategy was attempted in software for preventing the loss of approach control mentioned above, but the system too frequently recognized the one exception, "turn right heading 300," for missed approach. An improved version of this software reportedly was delivered with four days remaining in the evaluation period. The error is serious but eliminating it is feasible.

The PARTS instructors reported that the system seemed to fluctuate in speech recognition accuracy. The instructors stated that on certain days the computer speech recognition accuracy dropped considerably from its usual performance. The instructors referred to this phenomenon as the system "going dumb", and said that it seemed unrelated to variables suggested by Canyon, such as students, syllabus levels, time of day, etc. In systems as complex as PARTS, the many interactions between hardware and software could lead to such intermittent problems. Finding the sources of these reported lapses in recognition accuracy would be essential for an operational training system.

COMPUTER SPEECH SYNTHESIS IN PARTS

The VOTRAX speech synthesis or speech generation system was used in PARTS to simulate the speech of several different people who normally interact with the PAR controller. The pattern controller gives control of the aircraft to the PAR controller in the handoff procedure. The voice of the pattern controller during handoff is simulated by the speech synthesis system. At the end of the approach, the PAR controller may handoff control of the aircraft to tower or again to the pattern controller, in the case of a missed approach. Here again, the voice of the pattern controller is simulated through the speech synthesis system.

PILOT'S VOICE. The pilot verbally responds to the PAR controller's transmissions until "do not acknowledge further transmissions" is given, usually about five miles from touchdown. The voice of the pilot is simulated by the speech synthesis system, and provides good feedback to the student controller. For example, if the student gives a transmission "Wheels should be down, over", the pilot will respond "Roger, wheels down and locked". This audio feedback from the simulated pilot lets the student know that the system understood his wheels check transmission. Similarly, turns to final, or the initial vectors given by the student are given responses such as "Roger, turn right heading 165". This is excellent feedback for the student controller. It lets him know that the "pilot" understood his vector and is complying with the transmission.

The pilot's simulated responses also provide information to the student about the shortcomings of the speech recognition system. The student may ask for a right turn to heading 165 for example, and the pilot will respond with "Roger, turn right heading 145". This may be frustrating to the student, but at least he has some information about what the pilot understood. One difficulty here, is that the pilot's verbal response to the student does not always correspond with the aircraft's response. This is confusing to the student when, for example, the pilot repeats the proper vector, but the aircraft fails to turn.

The speech synthesis system was readily understandable. No students complained of any difficulty in understanding "Egor", the nickname given to the VOTRAX voice, and later used by students with reference to the entire system. Two minor problems with the speech synthesis system were differentiation between people and slowness of speech. Early in training, some students had difficulty differentiating between speakers at the beginning of an approach, when the same synthesized voice was used to represent both the pattern controller and pilot. If slightly different voices were used, this would ease the students' task slightly. However, with one day's experience students were readily able to discriminate the messages based on the context of the speech.

The slow pace of the synthesized speech occasionally caused difficulty in the timely execution of PAR procedures because the student had to wait for "Egor" to finish responding verbally before the next vector could be given.

MODEL CONTROLLER. Another use of the speech synthesis system in PARTS was as the model controller. The model controller concept used speech synthesis to simulate the actions of an ideal PAR controller. This is an excellent opportunity for the student to observe (visually and aurally) proper control procedures. A large part of this observation is listening to the synthesized speech of the model controller giving the proper transmissions at the proper time (relative to the PAR display). Slight difficulties were noted here in the lack of time delay between the voice of the model controller and the voice of the pilot, which were the same voice. For example, in the radio check procedure the model controller gives the transmission "how do you hear me", and the pilot responds "loud and clear". These transmissions were run together into one phrase that sounded something like "how do you hear me loud and clear". Some students were slightly confused by this the first one or two times they observed the model controller, but the PAR procedures are so standardized that there was no reported difficulty in identifying the speaker after brief practice. The major problem with the model controller was a lack of demonstrated skilled performance as a PAR controller. However, this problem was not due to the synthesized speech, and will be discussed later in this report (see "Demonstration Mode" in Section VIII).

MIMICKING OF SYNTHESIZED SPEECH. The Canyon observers agreed with the PARTS instructors that certain students tended to model their speech after Egor. This practice was specifically prohibited by the instructions, but some students seemed to be unable to refrain from mimicking the synthesized speech during voice training. Once this practice began during voice training, the students could achieve good speech recognition only by consistently speaking in that manner. Consesquently, these few students tended to speak somewhat like the synthesized voice throughout their training period. The PARTS instructors reported that the students who mimicked the synthesized voice tended to have problems with speech recognition.

In a related event, a discussion with one of the PAR instructors (not assigned to PARTS) revealed his misconception about one of the objectives of the system. He believed that the system was designed to teach students to speak like the synthesized voice. The instructor was relieved to find out that this was not the case, because Egor does not sound like an experienced Air Traffic Controller. Similar issues of user acceptance of innovative technology are discussed later in the report (see Section IX).

SECTION VI

TRANSFER OF TRAINING (TOT) STUDY

PURPOSE

The purpose of the transfer of training study was to assess the training effectiveness of PARTS relative to the traditional PAR training which takes place in classroom and PAR laboratory. This comparison of the two training methods was accomplished through a pseudo transfer of training test administered to students who had been trained in PARTS or the PAR lab. An ideal way to accomplish a TOT study is to test the student in an operational environment, but this was impossible in the present study. Therefore the pseudo TOT design was used, in which the controller's task at the air field was simulated by using PAR laboratory equipment. The transfer of training test was designed with two objectives in mind: 1) to be a job sample test, requiring the skills necessary to conduct a PAR approach at the field, and 2) to be a fair and unbiased test in which students from both training methods would have an equal chance of performing well.

METHODS

SUBJECTS AND APTITUDE GROUPING. The experimental plan called for two subjects from each training method, PARTS and PAR lab, to be tested each week for at least six weeks. Only two subjects could be trained on PARTS each week because of the single student console design of the prototype PARTS. An equal number of subjects in each training group was selected to represent three levels of prior grade point average (High, Medium, Low) in prior courses at the Air Traffic Control Schools. This stratified sampling procedure was used to enable assessment of subject by treatment interactions, i.e, to determine whether students with different aptitudes would benefit differentially from automated training (Chronbach and Snow, 1977). Cut-off scores for the aptitude groups were established by analysis of 100 recent ATC Schools grade point averages plotted as a frequency distribution. The sample was divided into thirds to establish the group criteria. This procedure established a two by three experimental design -- two methods of training and three levels of grade point average.

Two students were selected each week from the normal class size of 14 students to receive training on PARTS. Their participation was voluntary, but all students who were selected chose to participate. The selection was announced formally by a Senior Chief. Students reactions to this type of selection and to participation in automated training generally were quite positive. Information about student attitudes was obtained by questionnaire and interview techniques, and is reported in Section IX. On the final day of PAR training, two more students were selected from the remaining 12 to participate in the TOT study. Their selection was based on the same criterion of aptitude grouping according to prior ATC Schools grade point averages. Participation in the TOT study occurred on the final day of PAR training, after the students in both training methods had completed their final exam (P-run).

TRAINING METHOD. The two training methods represented in the TOT study were PARTS and normal PAR training. The normal PAR training course at the ATC Schools consists of approximately one half day of lecture, primarily on PAR phraseology and procedures, followed by a test the next day on phraseology. The remaining four days are spent in the PAR laboratory. Half the time, students in PAR laboratory practice controlling approaches, and half the time they act as bug operators, simulating the pilot responding to PAR controller transmissions. No changes in the normal PAR training course were instituted to accommodate TOT testing.

The PARTS students proceeded through the automated syllabus, but there was some orientation of the instructors toward the TOT testing. These modifications in instructional techniques may have been instituted not so much in response to the early TOT test results, per se, but because the instructors became aware of some of the shortcomings of the automated syllabus and the requirements for additional tutoring. The TOT study was the last procedure in the overall PARTS evaluation and the instructors' role had been evolving throughout this time. By the end of the study, the TOT testing, the instructors were taking a more active role with the system. The student who completed PARTS training during the TOT study was not merely a product of the automated system but of the combination of the automated system, personal tutoring from the instructor, and support from prior students. The support from prior students occurred through procedural changes resulting from another study carried out by Learning Designs, focusing on the personal interactions between instructors and students and the role of individuals in an automated training system (Joplin, 1980). It is the opinion of the Canyon team, supported by the PARTS instructors, that the students who completed the PARTS training at the end of this study were more proficient than some students early in the PARTS evaluation who were trained with the automated syllabus with very little instructor intervention.

TESTING PROCEDURES. The TOT test was developed by Canyon in conjunction with the PARTS instructors. The test consisted of three approaches, one from a left base, one from a right base and one straight in. One of the three approaches included a special situation, tower-clearance not received. All three approaches included full handoff procedures. The handoffs were given by an instructor. The equipment used for the approaches was the 15G19 PAR indicator, used in the normal PAR training course. This equipment was considered appropriate because it is similar to the PAR equipment used at the field. Different aircraft call signs were given on all three approaches, and no memory aid was available to the student to provide information about aircraft call signs or the type of approach requested (touch and go, full stop or low approach).

The TOT test was a change from the normal training procedures both in PARTS and in normal PAR. In PARTS, the small CRT constantly displayed the aircraft call sign and the type of approach requested. The student did not have to remember this information because it was available at any time during the approach, simply by glancing over at the small CRT. In the TOT criterion test, the information was not available, but, as in the real world of the air field, the student had to listen carefully to the handoff for this information, then remember it for proper use during the approach.

For the PAR students, the procedures of the TOT test also were somewhat different from normal training. PAR students generally did not practice handoffs during the training course. The aircraft call sign remained the same on all approaches during a session. The student did not practice remembering new call signs on each approach. In the TOT criterion test, the PAR student encountered a different aircraft call-sign on each approach.

Several other procedures or conditions included in the TOT criterion test were different from practice sessions in PAR lab because the test was designed to be representative of procedures used at the field. Another purpose of changing procedures was that PARTS students were at a disadvantage in transitioning to new equipment for the criterion test, while the PAR students were able to take the criterion test on the same equipments used in their training. It was felt that instituting some new procedures to be learned for the PAR students would decrease their initial equipment-familiarity advantage. Table 8 lists the characteristics of the TOT criterion test that differed from the normal training procedure.

A very brief transition training period was given to help all students learn the new procedures for the TOT criterion tests. Transition training was given by the PARTS instructors at the actual test consoles prior to administration of the TOT criterion tests. The transition training was approximately 1.5 hours long. Each student was able to control approximately three to seven approaches using the new procedures. The transition training improved over the 12 weeks of TOT testing. After the first few weeks, instructors were remaining with the students throughout the period of TOT transition training in order to simulate handoffs, as they would be given in the criterion test.

The TOT criterion testing was accomplished by three instructors. One acted as pattern controller, initiating the approach and giving the handoff to the student. Another instructor served as the simulated pilot ("bug operator") and the third instructor (from a more advanced portion of the ATC School curriculum) was the grader. He was uninformed as to the group membership (method of training and aptitude) of the students being tested. His job was to observe the student performance in PAR approach control, using a specially designed grading sheet to mark each action and transmission. The grader did not score the student's performance. The grading sheet included all information about a student's action and whether or not the action was correct or incorrect, but the grading sheet did not include scoring criteria. The instructor/grader was not informed after testing as to the training method from which the student came.

TABLE 8. TOT TESTING CONDITIONS AND DIFFERENCES FROM PAR AND PARTS TRAINING

TOT Test	Normal PAR Training	PARTS Training
1. Wind 140° at 10 or 180° at 10	Rarely used	Variable direction and speed, including gusts
2. Pilot Responsive PAR Instructor	Student "bug" Operator	5 Skill levels, variable assignment
3. Aircraft One type, 150 kts	Same as TOT	4 types, different speeds (fastest = 156 kts)
4. Communications Panel 30 button matrix	Same as TOT	8 button matrix
5. Handoffs Full handoff from pattern - controller	May not have used handoffs	Handoff on each practice approach
6. Aircraft Target Presentation on Scope	Same as TOT	Simulated target, enlarged with better edge definition

The TOT criterion test was administered to one student at a time. Each administration required the three instructors indicated above. Testing time was approximately 20 to 30 minutes per student. With rare exceptions, four students were tested each week, two from each training method. The students were not given feedback about their performance after the test session.

SCORING. Scoring criteria were developed primarily by PARTS instructors with guidance and assistance from Canyon. The objective of the scoring criterion development was to weight student errors in terms of their importance as indicators of proficency as a PAR controller.

The scoring criteria are given in Table 9. After the criteria for scoring were developed, each PARTS instructor scored all the grading sheets independently. The three approaches for each student, therefore, had two scores given. The inter-rater scoring reliability was extremely high, r=.92. The two scores on each approach were averaged and the three approaches for each student were averaged to provide a single score representing the student's performance on the TOT criterion test.

TABLE 9. TOT TEST SCORING CRITERIA

Penalty Points (Subtracted from 100 Points)	Student Controller Action
7	Failure to Handoff to Pattern Controller on Touch and Go
	Failure to give Wind and Clearance
5	Major Vector Error a. Allowing A/C to go from well Left/Right of course to well Right/Left of course b. Well Left/Right of course, no turn given c. "Stop Turn" given in normal PAR approach
4	Range Call - omitted or wrong time Touch and Go given "Contact Tower After Landing" Begin Descent Call - improper time Approach Glidepath - improper time Wheels Check - not given Course information incorrect or not given Medium Vector Error a. On course well Left/Right of course b. Left/Right of course to Right/Left of course c. Insufficient or excessive vectors d. Left/Right of course, no turn given
3	Call Sign incorrect or not used Handoff Procedure from pattern to PAR incorrect Over Landing Threshold omitted or not given at proper time
2	Minor Vector Error a. 1° turn given b. Request heading c. Early/Late Vector d. 360° turn

RESULTS

Table 10 shows the TOT scores of 29 students as a function of the two training methods and three levels of ATC School grade average. The table includes the marginal means. These data were subjected to a two by three analysis of variance. The analysis of variance summary is given in Table 11. The main effect of training method was not significant ($F=0.01$, $p>.05$). The effect of aptitude (grade average) was significant ($F=3.78$, $p<.05$) but the interaction of the training method and grade average variables was not significant ($F=0.37$, $p>.05$).

Since the Aptitude (prior ATC School grade average) variable was significant, a post hoc comparison among three means was done by the Newman-Keuls procedure. This analysis showed that the low aptitude group scored significantly lower on the TOT test than either the medium- or the high-aptitude groups ($q_{3,26} = 3.75$, $p<.05$; and $q_{2,26} = 2.99$, $p<.05$, respectively). No significant difference was found between the medium- and the high-aptitude groups ($q_{2,26} = 0.59$, $p>.05$).

These effects are plotted in Figure 4, where it is obvious that the low grade-average group scored lower than the medium and high groups on TOT test scores, but the effects were equivalent for the PARTS and PAR

TABLE 10. TOT TEST SCORES FOR 29 STUDENTS GROUPED BY TWO METHODS OF TRAINING AND THREE LEVELS OF APTITUDE (PRIOR ATC SCHOOL AVERAGE)

		Training Method		
		PARTS	PAR	
APTITUDE	High	1. 79*	1. 85	
		2. 88	2. 92	M=85.0
		3. 95	3. 77	SD=6.59
		4. 85	4. 76	N=9
		5. 88		
	Medium	6. 83	5. 92	
		7. 93	6. 93	
		8. 89	7. 83	
		9. 78	8. 84	M=86.7
		10. 86	9. 90	SD=4.75
			10. 83	N=12
			11. 86	
	Low	11. 65	12. 84	
		12. 72	13. 82	M=76.1
		13. 87	14. 51	SD=13.14
		14. 77	15. 91	N=8
		M=83.2	M=83.3	Grand M=83.2
		SD=8.23	SD=10.34	SD=9.22, N=29

*100 Points = Perfect Score

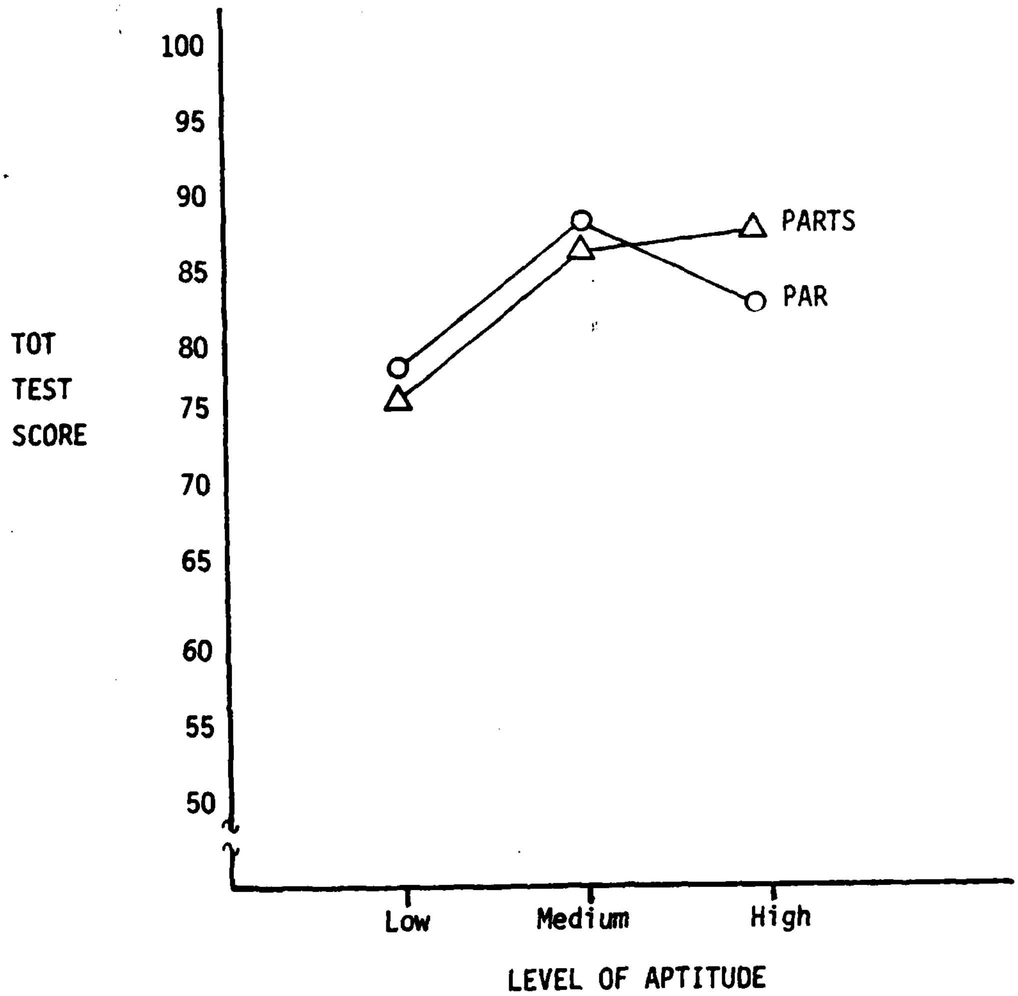


Figure 4. Mean TOT Test Score as a Function of Level of Aptitude (Prior ATC School Grade) for Two Training Methods (N=29).

TABLE 11. ANALYSIS OF VARIANCE SUMMARY OF TOT TEST SCORES

Source	SS	df	MS	F	P
Training Method	0.6	1	0.6	0.01	ns
Aptitude (Prior Grade Average)	574	2	287	3.78	.04
TMxPGA	57	2	28	0.37	ns
Residual	1748	23	76		

training methods. One seemingly strange result was that the high grade-average students from the PAR training course did not perform as well as might be expected by observation of the trends in the low and medium grade-average groups. The reason for this relatively poor performance is unknown, and may be simply a chance effect, since there were only four students in that group. As seen in the figure, the slight trend in the low and medium grade-average groups was for better performance after PAR training than after PARTS training. However, this trend was small, not statistically significant, and entirely erased in the overall comparison of the two training groups, by the relatively poor performance of the high grade-average PAR training group.

Supplementary data from the TOT study were provided by an overall performance rating given by the initial grader at the end of each approach. Each approach was rated on a five category scale, ranging from "Excellent" to "Inferior". Since each of the 29 students controlled three approaches, a total of 87 ratings was obtained. The percentage of approaches rated in each category is given in Table 12 for the two types of training methods. No distinctive differences between the PARTS and PAR training were apparent, particularly if the "Poor" and "Inferior" categories were combined. No approaches by a PARTS student were rated "Inferior", but sixteen percent of the approaches by PAR students were rated in that category. On the other hand, forty percent of the PARTS approaches were rated "poor", compared to only twenty percent of the PAR approaches.

TABLE 12. TOT APPROACH PERFORMANCE RATINGS FOR TWO TRAINING METHODS (IN PERCENT)

	Grader's Rating of Approach					
	Excellent	Good	Acceptable	Poor	Inferior	
PARTS	5%	24%	31%	40%	0%	N=42
Normal PAR	4%	31%	29%	20%	16%	N=45

DISCUSSION

The results of the transfer of training study indicated no difference in controller performance between the two methods of training. These results must be interpeted in light of additional considerations entering into the training procedure used in PARTS and the criterion test for the TOT study.

It may be argued that the PARTS students were not merely PARTS students but "PARTS/Instructor/Other" students. The PARTS instructors discovered early in the PARTS evaluation that students would be unable to complete the course in the five days allotted. Instructors stated that they became assertive in their use of the automated system, to push the student through the syllabus to complete the P-run within five days. During the period of the TOT study, the instructors were using override regularly, instructing the students not to use certain modes of the system, instructing the students to ignore certain instructional material presented on CAI or in the student guide, instructing the student not to use replay-with-errors, not to believe the error messages reported when replay-with-errors was selected, and not to believe the actions of the model controller as representing correct PAR controller performance. Therefore, the student completeing the PARTS P-run was not simply a product of PARTS training. The instructors had taken a much more active role in structuring, streamlining and individualizing the instruction. They were using the system as a tool for instructing the student. This was not the concept of the system, which was designed to lighten the load of the instructor by automating instructor functions. The change in instructor roles during the PARTS evaluation cannot be considered indicative of the application of automated voice technology to training, but of weaknesses in the PARTS courseware and instructional technology perceived by the instructors.

The PARTS students also were aided by former students who returned to assist in solving conceptual or practical problems with the system. Additionally, the two students assigned to the system during the week were scheduled for a brief (1/2 hour) overlap period during the day when they could exchange information and opinions about their experiences with PARTS. These increased person-to-person interactions were part of the study by Learning Designs. While the influence of these procedures is difficult to evaluate, it may have enhanced the ultimate learning and performance of the PARTS student (Joplin, 1980). In summary, the scores on the TOT criterion test may have reflected more than the instruction derived from the automated system. In this respect, the scores would tend to overestimate the training effectiveness of the prototype PARTS.

However, the TOT criterion test was unable to include the assessment of certain variables that were included in PARTS training. For example, wind variability was included in PARTS, as well as variability in pilot performance and different aircraft approach speeds. These variables are not included in normal PAR training nor were they able to be included in the TOT criterion test. In this respect, the scores on the TOT criterion test would tend to underestimate the proficiency of PARTS students in controlling a PAR approach.

In conclusion, the transfer of training data indicated that training on PARTS yielded proficiency equivalent to the normal PAR training course.

A note of caution, however, must be added here. The sample size of the present study, 29 students, was as large as possible given the time and equipment constraints. This sample size would be adequate only to reveal relatively large differences in the outcomes of the two training methods. Chronbach and Snow (1977, p. 46) suggest that studies attempting to look at aptitude by treatment interactions should employ large sample sizes, on the order of 100 subjects per treatment, to reduce sampling errors inherent in interaction effects. It remains possible that very large sample sizes could reveal the training method or the interaction in the TOT study to have a statistically significant effect. The present sample was large enough to indicate that such differences, if they exist, would be of small magnitude and of little practical consequence.

SECTION VII

PERFORMANCE MEASUREMENT VALIDATION STUDY

STUDY PURPOSE AND DESIGN

The Performance Measurement Validation Study was designed to determine whether the PARTS performance measurement system could discriminate between two groups which differed in PAR experience. The experimental plan called for obtaining PARTS P-run scores from 10 PAR qualified instructors and from 10 students with little or no background in PAR. This plan proved to be unworkable because a subject (student or instructor) could not be scored on the P-run without progressing through the entire PARTS syllabus. The time required on PARTS to reach the P-run was approximately 30 hours. It was not feasible to spend 30 hours per subject, particularly for the instructors, without interfering with the ongoing training program.

A short syllabus was designed jointly by Canyon and the PARTS instructors, with suggestions from a Logicon, Inc. representative. Final coding of the short syllabus was done by Logicon, Inc. The objective was to eliminate the PAR training portion of the syllabus as much as possible, and to reduce to a minimum, the time necessary for voice sampling and instruction on procedures specific to PARTS, such as the communication console. The number of practice approaches was minimized. The short syllabus included three approaches for the P-run, rather than the one approach found in the normal PARTS syllabus.

In addition to the PARTS P-run scores on three approaches, two instructor-observers graded each approach using the normal PAR lab scoring procedure and worksheet. Therefore, the experimental design enabled not only a comparison of PARTS P-run scores for qualified and inexperienced subjects, but also a comparison of PARTS scores with instructor-graded scores.

SUBJECTS

The number of subjects for this part of the evaluation was limited to four students and five instructors. The students varied in PAR experience from none (about to begin the PAR course) to one recent graduate of the PAR course. The five instructors were field-qualified PAR controllers with radar experience ranging from 2-20 years.

RESULTS AND DISCUSSION

The PARTS performance measurement scores and the instructor-graded scores are given in Table 13 for four students and five instructors, who each controlled three PARTS approaches. A data transformation was necessary on the instructor-graded scores because PARTS scores are based on 100 points per approach and PAR lab scores (instructor-grading) are based on 33 points per approach. Each instructor grade was multiplied by three to generate nearly equivalent scales for the two types of scoring.

PARTS does not give a single score for a P-run approach. Ideally, some weighted combination of the 15 skill categories would yield a summary score. A single PARTS score for each approach was generated simply by giving unit weights to the categories and calculating a mean PARTS score. While the utility of this mean score for training may be questionable, it is unbiased with respect to discriminating between the two groups (instructors and students). The two instructors' grades for each P-run were not scored by the same two instructors for all subjects; between and within instructors variability contributed to the scores, although all instructors followed the well-defined set of grading criteria established for the PAR lab.

The data in Table 13 show that the mean PARTS (automated) score for students was 87.9 compared to 93.4 for instructors, a difference of 5.5. The same comparison, based on the average of the two instructor-graded scores resulted in a mean P-run score of 34.6 for students and 82.2 for instructors, a difference of 47.6.

Based on the mean difference scores alone, the automated performance measures appeared to be considerably less effective than the instructor scoring in discriminating between the experienced and inexperienced controllers. However, mean difference scores alone do not provide sufficient information about group discrimination. The extensive range of instructors scoring (0-99) was due to the heavy penalties imposed for a "safety error", i.e. any controller error that could endanger the safety of the approach. This scoring procedure expanded the range of scores and contributed to the substantial mean difference between the experienced and unexperienced controller.

The three sets of scores, (PARTS, Instructor A, and Instructor B) on the 27 approaches were plotted in separate frequency distributions to portray the overlap of the scores for the experienced and inexperienced controllers. These three frequency distributions are shown in Figure 5.

Observation of the frequency distribution of PARTS scores in Figure 5 indicates a restricted range of scoring, with 26 of 27 scores falling within the range 81 to 100. This "ceiling effect" reduces disrimination between groups because a slight measurement error in the group of low scores could result in a greater overlap of the two distributions.

This "ceiling effect" problem can be attributed primarily to the high frequency of 100 scores given by PARTS. This is evident in Table 13, presented previously. For example, every score in performance category APT (azimuth position and trend) was 100. Therefore, this

TABLE 13. AUTOMATED RAW SCORES, MEAN, AND INSTRUCTOR SCORES FOR PARTS P-RUNS OF STUDENTS AND INSTRUCTORS

SUBJECT CONTROLLER	APPROACH NUMBER	HO	RCh	TTF	AGp	HA	APT	GPT	RCa	DH	CL	OLT	HR	EW	TB	TR	PARTS Mean	INSTRUCTOR SCORING A	INSTRUCTOR SCORING B
Student 1	1	100	100	90	80	93	100	100	100	100	50	100	80	100	100	87	92.0	54	63
	2	100	100	90	70	95	100	100	100	100	65	100	100	100	100	83	93.5	0	0
	3	100	100	70	70	83	100	85	100	100	50	100	70	100	100	93	88.1	63	45
Student 2	1	95	100	90	95	100	100	83	100	20	100	20	60	100	100	96	83.9	45	9
	2	85	90	90	85	97	100	95	100	100	100	100	80	100	100	100	94.8	72	63
	3	85	90	90	90	100	100	95	93	100	100	100	85	100	100	100	95.2	90	90
Student 3	1	80	90	88	80	100	100	93	70	100	50	100	90	25	100	100	84.4	63	45
	2	85	35	100	75	100	100	83	70	50	40	60	60	100	80	83	74.7	0	0
	3	100	70	94	100	83	100	50	68	100	65	100	65	100	100	100	86.3	45	21
Student 4	1	85	70	70	70	95	100	93	52	50	50	100	100	100	100	95	82.0	0	0
	2	100	100	54	85	100	100	75	100	100	95	60	100	100	100	73	89.5	0	0
	3	100	100	100	85	100	100	66	66	100	50	100	100	100	100	82	89.9	45	18
Instructor 1	1	85	100	100	100	100	100	100	100	100	100	100	80	100	60	89	94.3	99	81
	2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100.0	99	90
	3	100	100	90	95	81	100	100	100	100	100	100	100	100	100	100	97.7	90	72
Instructor 2	1	100	100	90	95	100	100	98	88	100	100	100	60	100	60	83	91.6	90	99
	2	100	80	90	95	100	100	96	100	100	100	100	90	100	100	100	96.7	90	99
	3	100	100	90	100	100	100	100	100	100	65	100	65	100	100	100	94.7	90	99
Instructor 3	1	65	70	100	85	100	100	85	70	50	100	60	60	100	100	81	81.7	99	54
	2	100	100	100	95	93	100	91	75	100	100	100	90	100	100	100	96.2	81	63
	3	100	100	70	95	90	100	68	80	100	100	100	70	100	100	100	91.5	0	63
Instructor 4	1	85	90	61	95	97	100	89	70	100	95	80	100	100	100	98	90.7	90	81
	2	100	90	90	80	100	100	100	85	100	100	100	80	100	100	100	95.0	90	81
	3	100	100	70	95	100	100	92	100	100	100	100	100	100	20	100	91.8	81	90
Instructor 5	1	100	100	90	85	86	100	89	100	100	95	80	80	100	100	96	93.4	72	72
	2	100	100	100	90	100	100	59	78	100	85	100	100	100	100	100	94.1	90	90
	3	100	100	100	100	100	100	89	80	100	100	80	60	100	80	94	92.2	90	81
MEAN SCORES:	Student Mean	92.9	87.1	85.5	82.1	95.5	100.0	84.8	84.9	85.0	86.7	86.7	82.5	93.8	98.3	91.0	87.9	39.8	29.5
	Instructor Mean	95.7	95.3	95.3	93.7	96.4	100.0	90.4	88.4	96.7	96.0	93.3	82.3	100.0	88.0	96.1	93.4	83.4	81.0

*Abbreviations of Skill Categories:

HO = Handoff
Rch = Radio check
TTF = Turn to final
AGp = Approaching glidepath
HA = Heading transmissions
APT = Azimuth position and trend
GPT = Glideslope position and trend
RCa = Range calls
DH = Decision height
CL = Clearance
OLT = Over landing threshold
HR = Handoff and rollout
EW = Emergency waveoff
TB = Transmission break
TR = Transmission rate

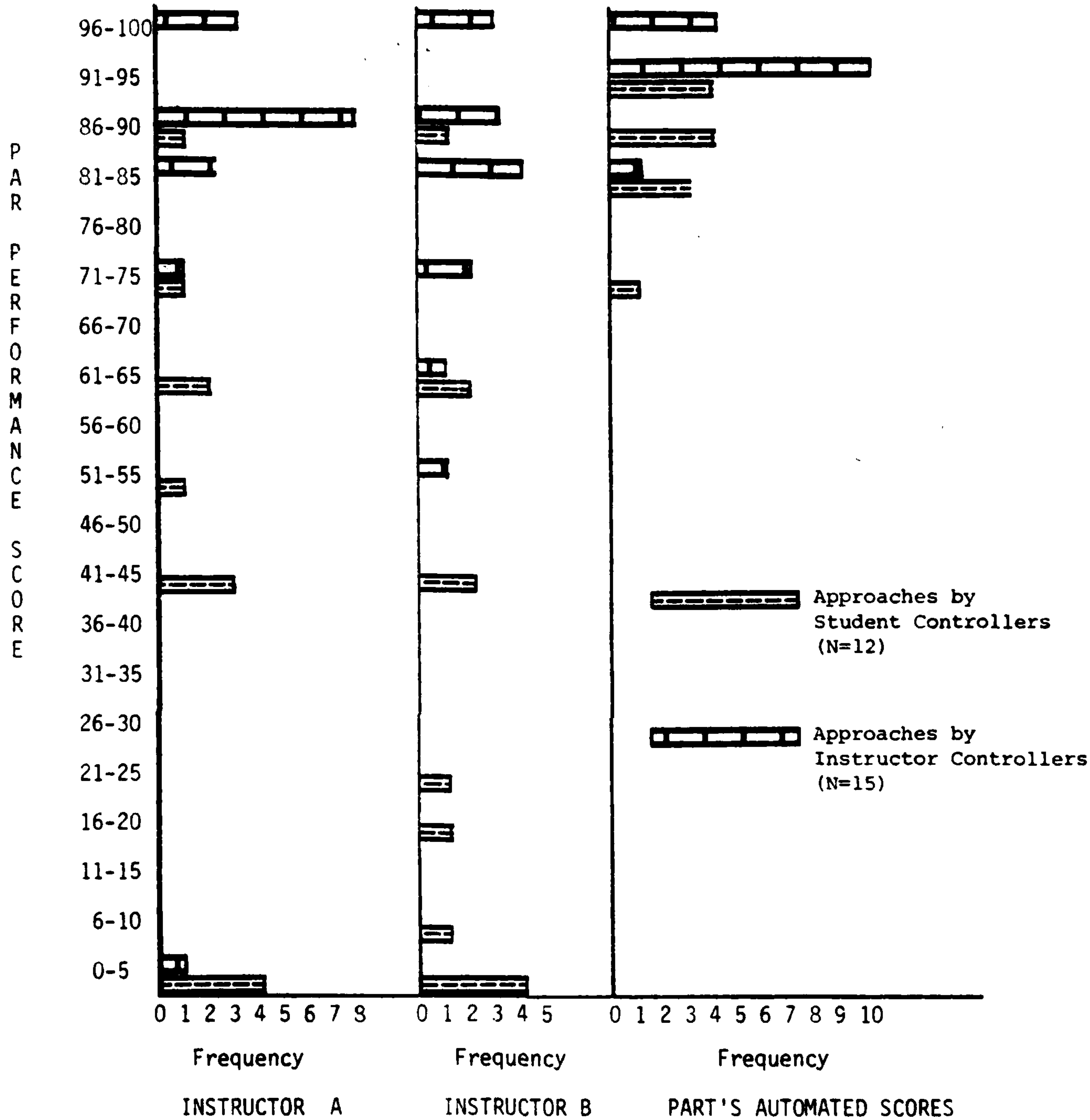


Figure 5. Frequency of PAR Performance Measurement Scores by Three Types of Scoring: Instructor A, Instructor B, and PARTS Automated Performance Measurement.

category provided no measurement power. It is very unlikely that the inexperienced students made no errors on azimuth position and trend. PARTS also assigned a score of 100 when no performance was observed in a category. This would contribute to the ceiling effect and to reduced discrimination between the skill levels of students.

An optimum cut-off score was determined for each of the three frequency distributions, based on a criterion of maximizing the percentage of correct group membership classifications. Table 14 includes the optimum cut-off score, the number of classification errors, and the percentage of correct classifications for the three sets of scores.

TABLE 14. DECISION CUT-OFF SCORES, GROUP CLASSIFICATION ERRORS, AND PERCENTAGE OF CORRECT GROUP CLASSIFICATION FOR AUTOMATED SCORING SYSTEM AND TWO SETS OF INSTRUCTOR SCORES

		Type of Scoring		
		PARTS Mean	Instructor A	Instructor B
	Optimum Cut-off Score	90.0	77	67
Classification Errors	Inexperienced (N=12 Approaches)	4	1	1
	Experienced (N=15 Approaches)	1	2	2
Percentage Correctly Classified	Inexperienced	67%	92%	92%
	Experienced	93%	87%	87%

Both sets of instructor scores (A and B) correctly classified 11 of the 12 approaches controlled by the inexperienced (student) group, and 13 of of the 15 approaches controlled by the experienced group. In terms of the percentage of correct classifications, both sets of instructor scores were ninety-two percent successful for the inexperienced group and eighty-seven percent successful for the experienced group. Corresponding statistics for the PARTS automated performance measurement system were sixty-seven percent correct for the inexperienced group and ninety-three percent for the experienced group.

In conclusion, the PARTS automated performance measurement system was moderately successful in discriminating between approaches controlled by experienced and inexperienced controllers. However, thirty-three percent of the inexperienced-controller approaches were misclassified. This type of bias in the measurement system was thought to be partially due to the assignment of perfect (100 point) scores on a default basis, when no performance occurred in that category.

In this study, the automated performance measurement system was tested for its capability to discriminate between two populations of controllers with vastly different experience in PAR. The system was moderately successful, but it's capability to discriminate between more subtle changes in performance that occur within one trainee as he increases in skill level, is still open to question. The accuracy of an automated performance measurement system is critically important when it is the basis for student feedback and for subsequent automated syllabus control. To successfully accomplish these functions, the automated performance measurement system must be capable of accurately measuring increments in a trainee's skill level as he progresses through training.

SECTION VIII

PARTS TRAINING FEATURES ANALYSIS

In this section a more detailed functional description of the experimental prototype PARTS is given. Evaluative comments based on observation and interview are included. System features other than computer speech recognition and synthesis are emphasised, but some overlap is unavoidable because of the interrelationship of the subsystems in PARTS.

INSTRUCTIONAL SEQUENCE

The PARTS syllabus is divided into seven levels: 1) introduction to the system; 2) azimuth control; 3) azimuth position, range, wind and clearance; 4) glidepath control; 5) approach termination, five-second rule, and landing threshold sequence; 6) practice and P-run, and 7) enrichment topics. A more detailed description of the syllabus is given in Table 15.

One of the stated strategies in the development of the syllabus was to build on the foundation of the student's prior knowledge. The first substantive PAR material in the syllabus essentially begins where the student ended in his previous course, ASR. Accepting the Handoff (from the pattern controller), Turns to Final, and Azimuth Corrections on Final, all deal with radio procedures and turn transmissions which are familiar to the student from ASR, although some of the details are slightly modified in PAR. The sequential introduction of new material is like an overlay technique, where the complete picture begins to emerge as succeeding layers are added. The student is continually responsible for all of the previous material plus the new topic. This approach to syllabus development seemed to work well in PARTS.

The sequential syllabus of PARTS was a distinct change from the all-at-once presentation of the material in the traditional PAR course. After one day of classroom lecture, the student controller in the PAR laboratory attempts to use all of the specified rules, procedures, and phraseology of PAR, even on the first approach. Errors and omissions decrease with practice and instructor feedback, until the student demonstrates mastery of the material on the P-run, after three days of PAR laboratory practice. Individual students may learn to apply different PAR procedures at different times during training. By contrast, the PARTS syllabus is considerably more structured and regimented. Concepts are introduced strictly in sequence.

The PARTS overlay syllabus is a sound concept, but improvements could be made. The instructional emphasis is uneven across levels of the syllabus. The first three levels are very detailed but subsequent ones seem to receive less emphasis. Glidepath control, Level 4, is a

TABLE 15. PARTS SYLLABUS

Syllabus Level	Syllabus Description
1.	The Precision Approach Radar Training System (PARTS)
	1.0 Introduction 1.1 Elements of PARTS 1.2 PARTS Syllabus 1.3 Using PARTS 1.4 Speech Recognition in PARTS 1.5 Use of INIT V/T 1.6 Getting Started
2.	Azimuth Control Procedures
	2.0 Introduction 2.1 Reviiew of Azimuth Radar Concepts 2.2 Checking Azimuth Alignment 2.3 Accepting the Handoff 2.4 Establishing Communications with the Pilot and Wheel Check 2.5 The Turn to Final 2.6 Azimuth Corrections on Final 2.7 Azimuth Control with Wind
3.	Azimuth Position and Trend, Range Information, Clearance Procedures
	3.0 Introduction 3.1 Course Position Information 3.2 Course Trend Information 3.3 Range to Touchdown 3.4 Clearance Procedure and Wind Information
4.	Elevation Control Procedures
	4.0 Introduction 4.1 Review of Elevation Radar Concepts 4.2 Checking Elevation Alignment 4.3 Approaching Glidepath 4.4 Do not Acknowledge Transmission 4.5 Begin Descent 4.6 Glidepath Position and Trend 4.7 Decision Height

TABLE 15. PARTS SYLLABUS (CONTINUED)

Syllabus Level	Syllabus Description
5.	Five second Rule and Landing Threshold Sequence
	5.0 Introduction 5.1 Five Second Rule 5.2 Over Landing Threshold 5.3 Rollout Instructions 5.4 Handoff to the Pattern Controller
6.	Practice and P-Run
7.	Enrichment Topics
	7.0 Introduction 7.1 Low Altitude Alert 7.2 Serving to Maintain Radar Contact 7.3 Emergency Waveoffs 7.4 No-Gyro Approach

particularly important concept in PAR, yet it receives noticeably weaker treatment than azimuth control, in Levels 2 and 3. PAR students have had previous experience with azimuth control in the ASR course, but they have not had previous experience in glidepath control. As one PARTS instructor succinctly put it, "We need less time in Level 2 and more time in Level 4."

In addition to the unequal instructional emphasis, the rationale for the sequence of concepts is unclear. The student doesn't encounter glidepath control until approximately halfway through the course. The students who trained on PARTS during this evaluation period controlled an average of 100 approaches in the six levels of the syllabus, through the P-run. Typically, forty percent of these approaches did not include glidepath. But glidepath and course (azimuth) control are the two primary control procedures to be learned in PAR. It would seem reasonable to introduce glidepath control immediately after azimuth control, and allow the student to practice this important dual-control responsibility prior to adding the requirement for range calls, tower clearance request and missed approach procedure. In the present syllabus these three procedures are taught before glidepath control. It is a difficult task to interpret the dual display and provide control for both azimuth and glidepath. PAR students need all the practice they can get on this dual-control task. Waiting for two days before introducing glidepath seems to reduce valuable dual-control practice time on the system.

INSTRUCTIONAL MEDIA

The media used for presenting information to the student consist of textbook, computer assisted instruction (CAI), computer generated speech, and simulated radar display.

The Student Guide textbook provides the student with basic information about PARTS, as well as rules, procedures, and phraseology of PAR. With minor exceptions, such as the topic of approaching glidepath, it is written clearly, illustrated well and aimed at a level consistent with the students' educational backgrounds. 181 pages, however, seems overly long for a supplement to an automated training system in a five day course.

CAI is used heavily in Syllabus Levels 1 and 2. It is effectively integrated with computer generated speech and digitized voice playback to serve as prompts for initial voice sample collection. The text material presented on the CRT is the primary method for introducing new information to the student. It is adequate, although too much reading from the CRT was required, particularly in light of the information overlap with the Student Guide. The best CAI periods involved frequent student interaction with the system, as in the section on azimuth alignment. Frequent interaction keeps the student's interest and allows him to develop confidence in this novel learning situation. PARTS instruction could be made more interesting to the student by using speech as a method for student/system interaction on the first day of training. Adding "Yes", "No" and "Next" to the phrases first trained on the system would enable the student to develop confidence by interacting with the system through speech. Currently the student communicates these three responses by pressing buttons on his keyboard.

CAI, as implemented in PARTS, represents an intermediate stage of instructional sophistication, between reading the Student Guide and interacting with the simulated PAR display during practice approaches. Some of the time devoted to CAI could be used more effectively, particularly in Level 2 of the syllabus. The CAI should bridge the student from the Guide to interaction with the PAR display, while keeping the reading material on the CRT to a minimum. PARTS is too sophisticated to spend much time "turning pages" for the student to read from the CRT.

The computer generated speech and the simulated PAR display are integral parts of the practice approaches which begin in syllabus Level 2 and continue throughout the remainder of the course. These media are used primarily as simulations rather than for instruction, and they are discussed elsewhere in the report. Brief demonstrations using the PAR display and the synthesized voice should be more frequent in the instructional phase of each syllabus level.

ADAPTATION AND REMEDIATION

Adaptive training has appeal in its promise to individualize instruction by changing the difficulty of the task as a function of the student's performance. Adaptive training may enhance learning by continually presenting tasks at an appropriate level of difficulty (Kelley, 1969). But, adaptive techniques, as Kelley (1969, p. 555) stated, "are by no means a panacea for the problems of training. ...adaptive training is new, and there are a hundred ways it can be done badly."

There are several ways to design adaptive training. Some of the most common are to vary the difficulty of a problem, the amount of time or practice on a set of problems, or to provide various routes through a syllabus via branching logic. Clearly, an automated system like PARTS has the capability to include all of these types of adaptive training. The experimental prototype PARTS included only an example of the possibilities of adaptive training by reducing problem difficulty when a student's performance scores fell below his average for the previous task. Also, a few extra practice problems are available at syllabus Level 6, Practice for the P-run. There is no syllabus branching logic designed to speed the superior student through the course or to provide the poor student with remediation or review. The PARTS syllabus could be considered mildly adaptive in some cases, such as Freeze and Feedback, where a task must be performed to a criterion. Some students can advance faster than others by achieving criterion performance sooner. The PARTS instructors found the syllabus pacing to be too slow, and they frequently exercised the override function to speed progress in certain areas.

The adaptive variables for reducing problem difficulty in the prototype PARTS were wind, pilot skill, and aircraft type. Reducing crosswind is a good way to decrease problem difficulty for students performing poorly on vectoring. Changing pilot skill levels was not an effective adaptive variable since the differences between skill levels were not great enough to be noticeable, with the possible exception of the worst (#5) pilot.

Assignment of a slower aircraft to a student having difficulty is a good concept, in theory. However, observation of PARTS indicated that the slowest aircraft, the U-21, was being assigned too often. Data extracted from the PARTS computer printouts by one of the instructors, ACC Cyr, indicated that of the four aircraft available, the U-21 was assigned on forty-four percent of the practice problems before the adaptive process was disabled. After eliminating the adaptive process, the U-21 was assigned on twenty-nine percent of the problems. During one observation session, a student was assigned the U-21 on five out of six approaches, although his controller performance appeared to be satisfactory. The U-21 may have been assigned as an adaptive variable because of errors in the speech recognition or performance measurement subsystems. The student was unhappy about seeing the slowest aircraft so regularly, since it slows the pace of the approach considerably. The total time required for one practice problem consisting of a U-21 approach followed by Replay with Errors was approximately 20 minutes. Because of the time delays induced by the frequent adaptive assignment of the U-21, the adaptive process was removed from the experimental prototype PARTS, slightly more than halfway through the evaluation period.

In summary, an automated training system such as PARTS has the capability to present instructional alternatives which are tailored to the performance of the individual student. Successful implementation of adaptive training techniques in any future version of PARTS is a difficult but worthwhile objective.

PAR PROCEDURES TAUGHT BY PARTS

In several instances, the procedures taught by PARTS differed from the procedures currently taught by the ATC Schools. Most of these differences were not critical. However, the PARTS instructors suggested that the instruction on "Approaching Glidepath" and "Begin Descent" procedures needs to be strengthened. Furthermore, some discrepancies exist for these procedures within PARTS subsystems - the Student Guide, CAI, and performance feedback. Students consistently had difficulty with Approaching Glidepath and Begin Descent on PARTS. Similarly, the instructors commented that PARTS instruction on Approach Termination (Over Landing Threshold) requires revision.

One major difference between the syllabi of PARTS and the ATC Schools PAR course is that the no-gyro procedure is included in the PAR course, and one of the three approaches in the P-run is a no-gyro approach. PARTS reserves the following topics for "enrichment", to be taught after the P-run: No-gyro, low altitude alert; servoing to maintain radar contact; and emergency waveoffs. Finally, the P-run consists of three approaches in the normal PAR course, but only one approach in PARTS. Although these discrepancies between PARTS instruction and current PAR policy and curriculum are not critical, they are important for two reasons. First, user acceptance suffers when the automated system teaches outdated or incorrect procedures. And second, the issue of flexibility of instruction in an automated training system is raised. As PAR procedures are modified over a period of years, would an operational automated training system be able to accommodate the changes without major revision and expense?

STUDENT/SYSTEM INTERACTION

STUDENT STATION. The physical layout of the PARTS student station is simple and efficient as shown in Figure 6. The student station consists of one large CRT used to simulate PAR approaches, a smaller CRT used for presenting instructional material and a communication box with multiple buttons for ICS and frequency selection. A head-set/microphone and foot switch microphone key also located are at the student station. These components are off-the-shelf items except for the communication (COMM) panel which was designed by Logicon. The COMM panel is the most complex unit at the student station but it has functional groupings of radio and ICS push buttons and is well designed and labeled.

The student station could be improved by putting all the components into a rack or, at a minimum, fix them to the table. Placing all the components of the student station into a rack also would reduce the problem of visual angle between the two CRT displays. During the PARTS evaluation the small CRT and keyboard were moved from the location shown in Figure 8 to the left of the large PAR display. Adjacent displays would require less physical motion on the part of the student to shift attention between displays.

INSTRUCTOR STATION. The PARTS instructor station components include a communications panel, printer, and CRT with keyboard, as shown in Figure 7. This equipment appeared to be well designed and located. No criticisms were voiced about the equipment at the instructors station or the system controller (computer and associated equipments).

The physical layout of the student and instructor stations are deceptively simple. This is a very complex training system but the complexity is not obvious from the physical layout. The complexities occur in the software and the technologies associated with computer speech recognition and synthesis, adaptive syllabus control, automated performance measurement, and simulations of the PAR display, pilot, aircraft, and wind.

STUDENT GUIDANCE. It was apparent from observing students that a considerable amount of effort had been put into organization of the various instructional methods such as CAI, demonstration, different modes of practice and so on. The student generally was led from one subtask to another with smooth transitions. However, this was not always the case, and some polishing remains to be done. Occassionally, students would get "stuck", particularly early in the course. Students were relying on prompts in the first day to direct them task by task. During this time their attention often was directed from the small CRT with visual prompts for voice training to the large PAR display for demonstrations, then back to the small CRT for more prompts to elicit speech samples. In some circumstances, prompts were seemingly omitted from the program and the student would be unsure what to do next. Since the instructors' station was in another room, at least 100 feet removed,

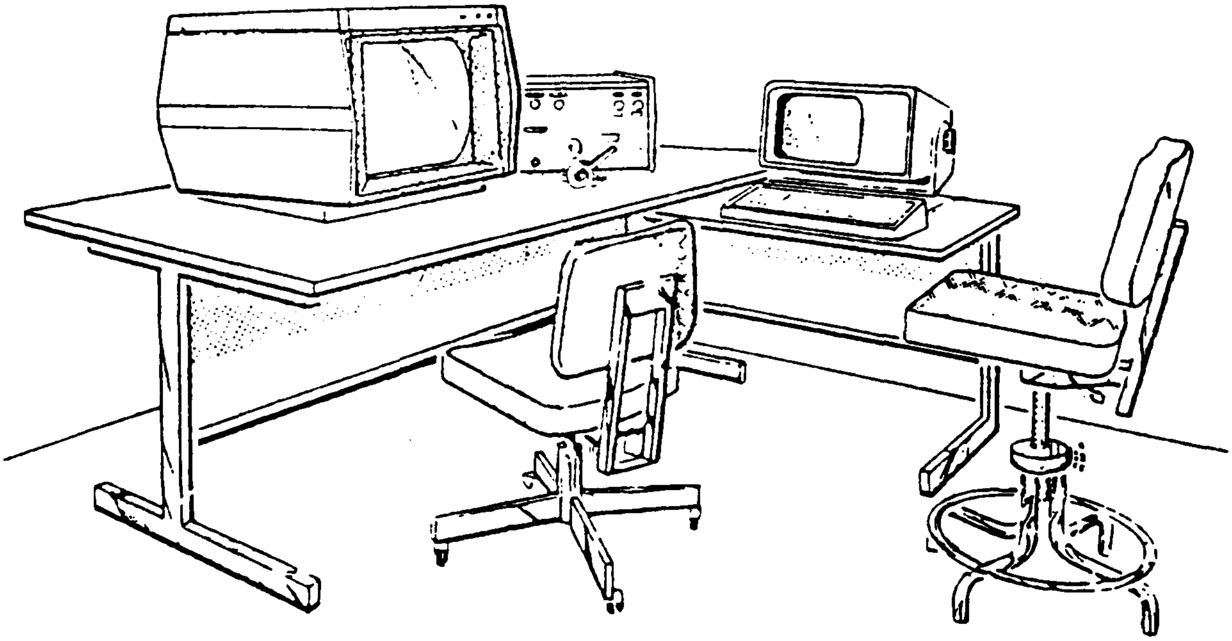


Figure 6. PARTS Student Station
(from Logicon, 1979)

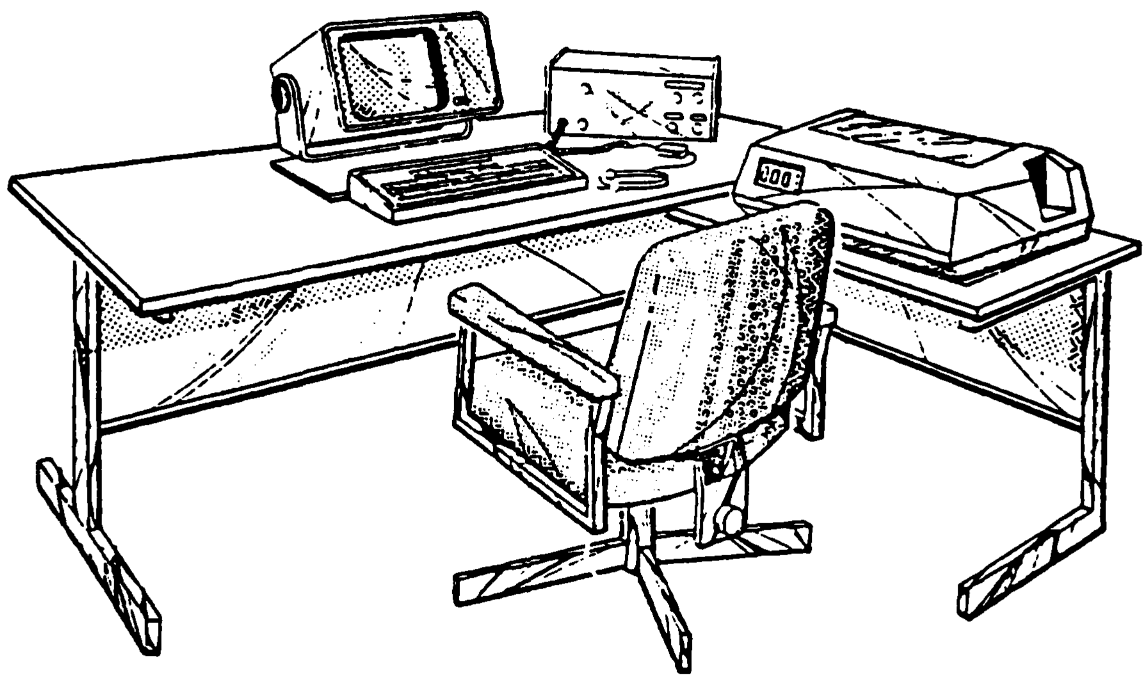


Figure 7. PARTS Instructor Station
(from Logicon, 1979)

students were reluctant to push the button for help, and would try to solve their dilemma by pressing the "next" and the "menu" keys. Smoother interaction between the student and the system could be obtained through further development.

In a related matter, one of the PARTS instructors suggested that more instruction was necessary regarding how to use the system. Some instruction is given to the student both in the student guide and in the CAI material. Syllabus level 1 provides a well organized, concise introduction to the system. Perhaps all that is needed is to give the student more frequent reminders on the proper use of the system. Particularly in the area of computer speech recognition, the student may need occasional reminders about items such as stylization, microphone placement and VU meter adjustment. A list of rules for good speech recognition is included on the inside cover of the student guide. This is an excellent idea, and could be put to further good use by occasionally including rules and reminders for the student as part of the courseware. Ideally, the frequency of such reminders would be keyed to the individual students success with the speech recognition system.

In general, many students seemed to feel somewhat lost during the first two days using the system. Toward the end of the training course, nearly all students were comfortable in working with the system. They knew what to expect, and how to effect smooth interaction with the system. More work needs to be done in easing the adjustment problem of the student during the first one or two days.

LOCUS OF CONTROL OF TRAINING OPTIONS. Locus of control is a concept applicable to the interaction between the student and the system, as well as the instructor and the system. The present use of this term is slightly different from the theory of personality construct by the same name (Phares, 1976). The term "locus of control" is used here to refer to the degree to which the student can make decisions about his own training. In PARTS, the locus of control is almost entirely in the courseware. Decision making by students is largely confined to within-task decisions regarding which radio transmission to give, i.e. performance as an air traffic controller. The student has virtually no decision making control over his own training, with two exceptions. One One was the capability to select replay with errors after each practice approach. However, this option was almost never used because of problems with the replay, which will be discussed later in this report. Another exception was in the option to select the freeze and feedback mode. Again, this selection was never used because of problems with that mode of operation (discussed in Section VIII under Modes of Operation). As a result of these two system problems, the student was left with virtually no options or decision making responsibility for his own training.

Automated training has the capability for great flexibility in syllabus branching, problem difficulty, selection of options, etc. Both students and instructors should have options available for guiding the training process, in lieu of automated guidance of the training. These comments are based on the observation that some students seemed to adopt an attitude of passive resignation after interacting with the system over a period of days. The role of the air traffic controller is to control the situation during a final approach. This means to actively make decisions and provide information necessary for safe landing. The nature of the automated training system placed the student in a very different type of role, namely being controlled. As the passive recipient of information, he was doing what he was told, moment by moment. It is well within the capability of automated training systems to provide the student with some degree of responsibility for the management of his own training. The student may desire to review certain information. He may desire to practice an approach with a particular aircraft, wind or level of pilot ability.

Students who are highly selected and motivated should be provided with the opportunity to excercise some training options such as selecting practice or review when they feel it necessary. This is not to suggest that the responsibility for these functions should be placed entirely on the student, but that some options should be available. For example, the student always has review and repeat options available in studying a text book. In the PARTS student guide, for example, he may decide to review a certain concept such as approaching glidepath. In order to review approaching glidepath he simply turns to the appropriate pages in the student guide. In a very small way, this is giving the student some responsibility for reviewing material when he chooses. The automated system could give the student the same option for reviewing CAI or for shaping practice approaches when the student feels that it would enhance his training. Students who desired to take a less active role in guiding their own training would simply default to the automated syllabus.

System design recommendations for shifting the training locus of control slightly toward the student are beyond the scope of this study. These comments are derived from observation of students on PARTS and from basic research in psychology (Broadbent, 1977; Powers, 1973). It is unknown whether implementing these concepts would have an effect on training effectiveness or user acceptance or both. However, we suspect that user acceptance would be enhanced, at least for most students.

SUMMARY. The interaction between students and system was generally good, although somewhat rough the first two days. Additional work appears to be needed to guide the student between tasks in the first few days of instruction. Later in training, when the student is more familiar with the entire system, he should be given the option to exercise some control over his own instructional alternatives. These concepts stem from the belief that students are not blank slates upon which the rules and procedures of PAR are to be written, but active, constructive individuals who are building schemata regarding the concepts, functions, duties and responsibilities of a PAR controller.

SYSTEM MODES OF OPERATION

Several modes of operation were included in PARTS, specifically, the Instructional Mode (Phase 1), Freeze and Feedback Mode (Phase 2), Practice Mode (Phase 3), and P-run Mode (Phase 4). The terms "Mode" and "Phase" seem to be used interchangeably in the PARTS documentation. The Demonstration Controller Mode and the Replay Mode are sometimes listed as distinct modes (Barber, et al., 1979) and sometimes subsumed under Phases 1 and 3, respectively (Hicklin, et al., 1979).

INSTRUCTIONAL MODE (PHASE 1). The Instructional Mode also is known as interactive teaching, and voice data collection. As the name implies, this mode is dedicated to teaching new PAR concepts and collecting voice samples for the relevant new phrases. The student is exposed to this mode almost every time he advances to a new level of instruction in the PARTS syllabus. The Instructional Mode is particularly evident in the first two levels. It reviews materials given in the Student Guide, gives demonstrations of new concepts and collects voice reference patterns.

The instruction provided by PARTS generally is good, but there are some problem areas, particularly the Approaching Glidepath, Begin Descent, and Landing Threshold Sequence procedures.

FREEZE AND FEEDBACK MODE (PHASE 2). Freeze and Feedback is an optional mode, in which the student can practice newly learned material and receive immediate feedback on errors. The practice problem is frozen only if a mistake is made on the new instructional material and error feedback is given on the CRT.

Limited opportunity existed to observe Freeze and Feedback because the students were instructed not to select this mode. This decision by the PARTS instructors is a meaningful critique of the way the Freeze and Feedback Mode was implemented in PARTS. They commented that it is a good concept, but unworkable in its present form because of the length of time required to practice a new PAR rule or procedure and to receive feedback on it. The problem is particularly evident for procedures that occur late in the approach, such as "At decision height". The student must work through four or five minutes of the approach before arriving at decision height. If the student makes an error on the Decision Height call, the problem is frozen and he receives feedback immediately. Ideally, he then would be allowed to correct his error promptly. But in the present design of Freeze and Feedback, the student again begins control of the approach from the start, and five or six minutes (and much cognitive processing) will intervene before he attempts to make the correct Decision Height call.

The PARTS instructors commented that once the optional Freeze and Feedback Mode has been selected, the student can't get out of it. If errors are detected, even invalid ones due to speech misrecognitions or failure of the performance measurement system, the approach is started over entirely. The student gets no "credit" for having completed an approach and therefore he can't progress to the Graded Practice Mode. The instructors suggested that an override selection should be available to allow the student to escape the Freeze and Feedback Mode, if he feels that training is being slowed down excessively.

GRADED PRACTICE MODE (PHASE 3). The Graded Practice Mode is a strength in the system. Yet it is the area in which improvements would be most beneficial. A good description of the Graded Practice Mode was given by Logicon, Inc. (Hicklin, et al, 1979, pages 24-28).

> Simulated control situations are presented and the trainee practices whatever skills he or she has acquired at that point in the syllabus. After the approach, feedback is given about the trainees performance on the new and previously learned material. A replay of the problem is then offered. The trainee can choose to observe the replay with or without the error reporting feature. The replay recreates all aspects of the approach. An actual recording of the trainee's voice is played back in sync with the aircraft dynamics. If the error reporting feature was selected, the replay stops when an error is detected and the error is explained on the CRT.

The Graded Practice Mode integrates all of the sophisticated technology of PARTS into a continuous simulated approach. The APE model, speech recognition, performance measurement, performance feedback and adaptive problem selection all are involved directly or indirectly. The utilization procedures at the ATC Schools (omitting Phase 2, Freeze and Feedback) meant that Phase 3, Graded Practice, was the primary mode used for instruction. Student controllers gained skill and proficiency in conducting PAR approaches primarily through practice in Phase 3. Appropriately, they spent the majority of time in that phase.

PERFORMANCE TEST (PHASE 4). The Performance Test (P-run) is the same as a Phase 3 Practice Approach, with two extra features. A special report of the student's performance on the P-run is printed out and the replay files are saved. These are valuable features for the instructor because they serve as a permanent record of the student's "final-exam" performance.

Currently, the PARTS P-run consists of only one approach. In the normal PAR training program, the P-run consists of three approaches. Multiple approaches are beneficial because they provide greater reliability and accuracy of performance measurement. The PARTS instructors recommended that the P-run be expanded to three approaches. On another occasion, the need for any P-run was questioned. Continuous, accurate performance measurement would preclude the need for a P-run.

The P-run performance measurement can be rescored to correct speech recognition errors. The P-run printout of recognized phrases can be compared to the actual (digitized) phrases during observation of a replay. The instructor notes any misrecognitions and follows the "Modify" procedure to correct them. Then a new set of scores is computed for the P-run. The Modify procedure is a necessary safeguard for the student. It ensures that his grade on the final exam is not reduced by speech recognition errors.

The PARTS instructors made one suggestion for improvement of the Modify procedure. Currently, a series of unrecognized phrases cannot be inserted. Only one unrecognized phrase can be entered. The instructors recommended that the Modify procedure allow insertion of all unrecognized phrases to ensure that the automated scores fully reflect the student's performance.

Observed PARTS P-runs were indicative of the strengths and current weaknesses of the system. The APE model, CSR, and simulated PAR display functioned adequately to enable the controller to conduct a successful PAR approach. As a final exam, however, the P-run fell short of providing accurate measurement and feedback. Speech recognition errors were the source of some of the problems in the P-run performance measurement, as discussed in Section V of this report.

The Modify routine was designed to allow the instructor to correct speech recognition errors after the fact, and allow the automated scoring system to rescore the P-run. In one P-run example, correcting nine major recognition failures by the Modify routine, resulted in the elimination of nine error messages, but nine new error messages were generated. Of the total 21 error messages before Modify, the PARTS instructors judged 11 of them to be invalid.

In summary, the P-run demonstrated the capability of PARTS to fully simulate a PAR approach, including aircraft, pilot and wind variables, to automatically measure and store records of student performance, and to record the entire approach for later replay. The performance measurement system requires debugging and further development.

DEMONSTRATION MODE. The four modes previously discussed have been termed "Phases of Instruction" in some of the PARTS documentation. The Demonstration Mode is a mode of operation of PARTS, but not a phase of instruction. The Demonstration Mode occurs in both Interactive Teaching and Voice Data Collection (Phase I), and when the system is operating but no student is signed-on.

The Demonstration Mode consists of an approach problem in which an automated Model Controller, rather than the student, conducts the approach. The Model Controller introduces himself to the student in Syllabus Level 1 with the following display on the CRT.

> Hi
>
> I'm your model GCA controller. Whenever you sign on you may notice a GCA approach in progress. That's me doing the ace controlling. By the end of this tutorial course you will be as good or better than I.
>
> Now don't let the pilot and pattern (ASR) controller confuse you. We all use the same voice. As you progress through training, you will learn to identify pilot responses and pattern controller advisories as well as the GCA PAR vocabulary.
>
> Watch this approach.

The concept of the Model Controller is excellent. His "voice" is clear and understandable, albeit synthetic. Psychological research has clearly shown that learning can occur through observing a "model" engaged in the desired behavior. *Ideally, the model should perform* without error, and demonstrate to the student highly proficient performance. The Model Controller is described as "a sophisticated simulation of an ideal PAR controller" (Barber, et al., 1979).

The PARTS Model Controller, when observed at the ATC Schools, was less proficient than desirable. Its performance on several approaches was scored by two instructors using normal PAR lab scoring method. It failed the P-run.

In the PAR laboratory procedure, a P-run consists of three approaches, each worth 33 points, for a total of 99. A passing score is 75 or above. The following are examples of errors made by the Model Controller while being given a P-run grade by the instructors.

A. First Approach

1. "Decision Height" was called early. This is a Safety Error as defined by the PAR laboratory scoring criteria (- 25 points).

2. "One mile from touchdown" called early (-3 points).

3. Missed one trend advisory and three glide-path position errors (-3 points).

   Score for approach #1 = 2 points

B. Second Approach

1. Called "Approaching Glidepath" late (-3 points).

2. No glidepath call given for 1/2 mile after "Begin Descent" (-3 points).

3. "On Course" called rather than "slightly left of course" (-3 points).

4. Two trend errors and two glidepath position errors (-3 points).

5. Not enough glide path calls (too many course calls) (-3 points).

   Score for approach #2 = 18 points

C. Third Approach

1. Five trend and nine glidepath errors (mostly calling "On glidepath" when A/C was "Slightly below glidepath") (-12 points).

2. Not a proper ratio of glidepath to course calls (3:1 is recommended). Not enough glidepath calls were given (-3 points).

   Score for approach #3 = 18 points

Total score for P-run = 38.

Passing score for P-run = 75 or above.

General comments of the PARTS instructors regarding the Model Controller's performance were:

1. Too many course calls were given. Needs to work toward the 3:1 ratio (glidepath to course call).

2. Glidepath position parameters need work.

3. Decision Height timing is too loose. (Note: this item was adjusted later in the PARTS evaluation period).

In summary, the Model Controller is a good concept for instruction. It could be used more often, particularly for demonstrating particular procedures rather than entire approaches. The Model Controller should demonstrate a high degree of skill. Failing a standard P-run is unacceptable performance. Program enhancement for the Model Controller is recommended.

"Canned" examples of approaches could suffice for instructional purposes, rather than a full controller model. Canned replays of ideal controller performance would be better than a fully flexible model controller that is error prone.

REPLAY MODE. Replay capability is available for Phase 3, Graded Practice, and for Phase 4, the P-run. The student can select replay with or without error messages. When "Replay with Errors" is selected the replay stops when a student error is detected, and an error message is displayed. After reading the error message, the student presses the "Next" button to continue the replay.

Replay with error messages is a good teaching technique. It gives the student specific feedback information on his performance as it occurs during replay. It avoids the disruptive effect of stopping the original approach to give error feedback. In Replay, the error message is given as soon as an error occurs, yet it does not disrupt the student control because the stopped action occurs during replay rather than during the original approach. "Replay with Errors" is a key instructional procedure in PARTS for guiding the student to develop PAR control skills. The major problems are the accuracy of the error messages (discussed later in this section) and the lengthly time required for replay with error messages.

PARTS does not provide the capability for the student to observe only a portion of the approach during replay. When the student selects "Replay with Errors" it may take up to 8 minutes for the replay, but the student may be interested in obtaining error feedback information on only a small portion of the approach. The time required for replay seems particularly long for the slowest aircraft, the U-21. As one student said "I wish Replay had a fast-forward". Because "Replay with Errors" is important for learning, students should not be discouraged from selecting it by the requirement to endure many minutes of delay. A capability to replay only selected portions of the approach would be an improvement. For example, if the student wanted replay of "Decision Height" he could select replay to begin at 2 or 3 miles from touchdown, rather than 9 miles from touchdown.

A discrepancy noted in observation of Replay was lack of synchrony between the aircraft on the PAR display and the reproduction of the students voice. This problem was detected during a discussion about the need for the student to make certain calls within very narrow time limits, such as "At Decision Height". The student made the call correctly in the original approach, but the call appeared to be late during the replay. Adjustment of this audio/video time discrepancy is important because the timing of some calls is critical. Instructors' attempts to evaluate a student's performance by observing and grading a Replay will be inaccurate when the audio/video asynchrony occurs. The estimated magnitude of the error was approximately one-half second, and it was not clear whether the problem was constant or sporadic. Attempts to fix the asynchrony problem reportedly were not entirely successful.

SIMULATIONS

SIMULATED PAR DISPLAY. The simulated PAR display is the primary information source for the student controller during a practice approach. The two main elements of the PAR display are the elevation (glideslope) and azimuth (course), as shown previously in Figure 1.

Generally, the simulated PAR display was excellent. The PARTS instructors made some suggestions for improvements, such as: representing radar "ground return" or the runway centerline reflector to provide a reference for the "over landing threshold" transmission; limiting the excessive range of cursor servo movement; checking early appearance of aircraft target on the glidepath display; and sometimes showing more than one aircraft on the display. The radar sweep in a normal PAR display was not incorporated in PARTS, but this omission is not expected to detract from training effectiveness since it is not used as a cue in controlling. A more beneficial addition, perhaps limited to an "enrichment" topic, would be to include a great deal of "clutter" and ground return, as found in PAR at the field. This visual "noise" changes the perceptual requirements and increases the difficulty of determining glidepath and course position and trend. Despite leaving some room for improvement, the simulated display was more than adequate for basic PAR training.

One capability of the PARTS display, not found in the normal 15G19 PAR display, is that individual components can be shown. In PARTS Syllabus Level 2, for example, azimuth procedures are taught by displaying only the azimuth portion of the display. This technique appears to be advantageous early in training because it eliminates extraneous information and helps the student focus on one part of the total task.

Another good feature of the PARTS display is the capability to summarize an approach by displaying the track history of the aircraft. This provides a rapid overview of the previous approach, and serves as a memory aid and information feedback to the student.

In summary, the simulated PAR display is well designed, realistic, and provides added instructional capability compared to the standard PAR laboratory display. The instructors and students uniformly commented favorably about the display, with a few minor exceptions and suggested improvements.

AIRCRAFT/PILOT/ENVIRONMENT (APE) SIMULATION. The APE simulation enables realistic movement of the aircraft during the approach. As the name implies, it models aircraft dynamics, pilot behavior, and wind characteristics to produce the aircraft target and movement on the simulated PAR display.

Four aircraft are included in the system, as shown in list below, which was adapted from Hicklin et al., (1979b).

Call Sign	Aircraft Type	Approach Speed (Kts.)
Army 876	U-21	98
Marine 687	A-6	115
Navy 310	P-3	130
Air Force 307	T-38	156

The use of four aircraft types in PARTS was beneficial for training. Students were able to experience the effect of aircraft approach speed on the pacing of PAR transmissions as well as the interaction of wind speed and aircraft speed. Another benefit of the multiple aircraft types was the use of different call signs. In the normal PAR training, by comparison, only one aircraft speed (150 kts) is used and, for an entire session (approximately 50 minutes), a student uses the same aircraft call sign. However, the assignment of the four types of aircraft in PARTS practice approaches needs improvement. It was reportedly pseudo-random, but excessive use of the U-21 aircraft occurred because of the problems with the adaptive process, as previously discussed.

Suggestions for improvement in the aircraft simulations were given by the PARTS instructors. They felt that the aircraft target "pushed over" too quickly on the "Begin Descent" transmission, and similarly, responded too sharply on turns. One instructor commented that the target appeared too thin, and too long. But the extra length was purposefully designed to ease the student's task of discriminating target subsections representing "On Glidepath" and "Slightly Above/Below Glidepath."

The pilot model included five levels of pilot skill. The good pilot usually was assigned early in training. Later, less skilled pilots were assigned to increase problem difficulty. Pilot skill level, like aircraft type, was an adaptive variable. The best pilot usually was assigned when a student was not performing well on azimuth or glidepath position and trend.

Unlike aircraft types, pilot skill levels were very difficult to detect while observing PARTS. The worst pilot may be detectable by wander from assigned heading, but he is very rarely assigned. One of the variables included in the pilot skill level is the percentage of glidepath transmissions ignored. This may not be a good variable to include because the student often interprets a failure to respond as a speech recognition error, rather than a pilot variable.

Review of the documentation of the pilot skill model indicated a relatively sophisticated and subtle design. It simulates the pilot's ability to infer actual glideslope, to correct for perceived errors in glideslope position and course, to respond with over- or under-corrections, and to engage in a "random walk" wander between correction advisories (Barber, et al., 1979). While the level of sophistication is admirable, in practice, the pilot model may be over designed in this application because its subtleties are overwhelmed by the effects of speech recognition errors. Differences between pilots responding to a controller transmission calling for a 5 degree right turn have much less effect on the approach than when the "pilot" (speech understanding system) interprets the transmission as a call for a 25 degree right turn. Speech misrecognitions are sufficiently frequent to mask pilot model differences. PARTS probably could provide good training with only one good pilot. A few approaches with a very bad pilot could be instructive at the end of the course as an enrichment topic.

The purpose of the environment simulation is to include the effects of wind on the approaching aircraft. A digital readout of wind direction and velocity is available to the student on the simulated PAR display. The beginning PARTS student quickly realizes that an aircraft may have to be assigned a heading of 162 or 164, for example, in order to maintain a course parallel to the 160 extended runway centerline. The wind information displayed to the student becomes meaningful when it gains functional significance for turns and vectoring. The simulated effects of different winds on different aircraft types is of particular instructional value. The wind model, like the pilot model, seems almost too elaborate for introductory PAR training. It includes a steady component, random component, and gusts (Barber, et al., 1979). Occasional errors in the model should be eliminated, such as the indicated wind moving the aircraft in the wrong direction. Overall, the wind model is a definite contribution to training, not found in the present PAR laboratory.

PERFORMANCE FEEDBACK

Whether one subscribes to principles of learning theory or a control systems viewpoint (Powers, 1973), the same concept is fundamental to learning and improved performance. That concept is known by several names: knowledge of results; information feedback; or error information. A person requires information about errors, (discrepancies between actual and ideal performance) for modifying his behavior on a subsequent trial. This is a fundamental precept of learning. Similarly, information about successful compliance of actual and ideal performance can serve as a reward to motivate further learning. The PARTS design was intended to provide several different levels of feedback to the student, including approach track history, grades by performance category, replay, and error messages.

When the PARTS student has completed a practice problem, several types of feedback are given, namely, track history, grades by performance category, and replay with error messages.

TRACK HISTORY. The first type of feedback given to the student after practice problem is a track history of the aircraft approach, which appears on the simulated PAR display. This is good, rapid feedback. The student can review at a glance the glidepath and course positions and trends because the entire track history of the approach is displayed at once. The track history remains on the display while the system computes the student's performance scores and grades for the approach.

GRADES BY CATEGORY. The next level of feedback given to the student is a listing of grades by category. The list is displayed automatically (without selection) on the PAR display. An example of a grade list is shown in Table 16, adapted from Hicklin, et al. (1979b).

Several weaknesses of the grading system were noted. Invalid "errors" were not uncommon, and the automated scoring system translated them into feedback of Satisfactory or Needs Work, even though the student made no PAR errors. After working with the system for a day or two, the student tended to disregard a Needs Work on a category in which he was satisfied with his performance. He will believe his performance was acceptable, and attribute the low grade (Needs Work) to a system error. This attribution may or may not be correct. To the extent that he is correct, the student more than doubles his training time efficiency by going on to the next problem rather than selecting "Replay with Errors". To the extent that he is incorrect, he will fail to detect and modify his unacceptable performance. The student is "second guessing" the speech recognition and performance measurement systems by assuming that a Needs Work or Safety Error grade was caused by system problems.

TABLE 16. EXAMPLE OF FEEDBACK ON PERFORMANCE GRADES

Skill Category	Grade
Accepting handoff	Perfect
Radio check	Perfect
Turn-to-final	Needs work
Approach glidepath	Satisfactory
Heading transmissions	Needs work
Azimuth position and trend	Satisfactory
Range calls	Perfect
Clearance	SAFETY ERROR
Handoff and rollout	Satisfactory
Transmission break	Perfect

During the evaluation period, the SAFETY ERROR grade was added, at the request of the PARTS instructors, to emphasize important PAR controller errors that endangered the safety of flight. The SAFETY ERROR flashed on and off to emphasize its importance. This concept was an excellent addition to the PARTS grading and feedback process. Unfortunately, the system frequently would give an invalid SAFETY ERROR for certain performance categories, such as Clearance. These frequent invalid SAFETY ERROR grades detracted from the meaningfulness of actual safety errors.

The student saw many Needs Work and SAFETY ERRORS, and they ceased to act as a signal to select "Replay with Errors." Even the PARTS instructors, familiar with the system, commonly received several Needs Work grades or a SAFETY ERROR message during a practice problem. In this case, the instructor, with some confidence, could attribute the grade to a failure in system. Students, by definition, cannot (and should not) be burdened with the responsibility for correctly attributing a low grade to either their own performance or the system.

The reward effect of an honest Perfect grade is diminished by giving Perfects for all performance categories in which no performance occurred. For example, if an aircraft is lost off the scope during an early portion of the approach, the remaining performance categories, such as Tower Clearance, Decision Height, and Over Landing Threshold, will receive grades of Perfect. The grading system begins with 100 points (Perfect) in each category, then subtracts points as errors are made. Where no performance has occurred, no errors have been made, and a Perfect score usually is obtained. A better procedure would be to give grades on each problem only for categories in which some student performance has been sampled and judged correct.

REPLAY WITH ERROR MESSAGES. The motivation for a student to select Replay with Error Messages is to determine what type of error he made. The motivation not to select Replay with Error Messages is the time required to gain the desired information feedback. For example, suppose a student is unsure whether he made a timing error in the "At Decision Height" transmission. With the slowest aircraft, the U-21, if he selects the Replay/Errors option, he must wait approximately 6-8 minutes (depending on the number of other error messages) to determine whether he made the error. Even selection of Replay without Error Messages would require about 6 minutes. The student is reluctant to wait that long. The strong tendency is to refuse the Replay with Errors option, go on to the next problem, and try to perform "Decision Height" correctly. In so doing, the student is an active, interested participant in the problem itself, compared to a passive observer, waiting for the desired feedback during Replay. Toward the end of the evaluation period, one PARTS instructor was telling his students not to select Replay with Errors. The other instructor found that students avoided it on their own, after a few tries.

DISCREPANCIES AND SUGGESTIONS FOR IMPROVEMENT. PARTS students did not obtain sufficient performance feedback because of a combination of two factors: 1) the four descriptor grades, given after each practice problem, lacked sufficient detail to enable the student to improve on a subsequent trial (in fact, the descriptors were not designed for that purpose) and; 2) a replay with error messages was unacceptably long, sometimes as much as 8 minutes. One possible improvement would be simply to list the error messages after each problem, either exhaustively or by category. In either case, the student could ascertain quickly whether speech misrecognitions or performance errors were involved. The availability of listings by category would allow the student to select the error messages only for particular categories, presumably those in which he received a Needs Work or Safety Error. For example, numbers could be added to the display, and the student could be given the option to see the error messages for any numbered category he selected. An example of this display modification is shown below.

	Skill Category	Grade
1.	Accepting Handoff	Perfect
2.	Radio Check	Satisfactory
3.	Turn to Final	Needs Work
4.	Approaching Glidepath	Satisfactory

SELECT ERROR MESSAGES FOR CATEGORIES (__,__,...)
OR NEXT

Compare this format to the unnumbered format in Table 16, presented previously. In this example, the student might choose to view the error messages for category #3, Turn to Final, because he received a Needs Work grade. Or he could view categories 2-4 where he scored less than Perfect. If, upon reading the error messages, he remained uncertain of his errors, then he could select Replay with Errors. This procedure would give the student an immediate overview of error messages. It would involve the addition of another level of detail of performance feedback information, between the descriptor grades and the lengthy Replay with Errors.

As suggested elsewhere in this report, a selectable range of replay would save much valuable time. If the student wanted to see Replay with Errors only for Decision Height, for example, he could select Replay to start at two or three miles from touchdown rather than nine miles. Presently, selection of Replay with Errors always starts at nine miles and stops for each error message in all categories.

A peculiarity of the PARTS performance feedback system was observed in error messages such as, "You were understood to say 'On Glidepath.' The correct position call should have been 'On Glidepath'." This nonsensical feedback is not helpful to the student.

Identification of speech recognition errors plays an important role in interpreting the validity of the error feedback messages. As an aid to the student, it would be helpful if the recognized phrases were shown on the display during replay.

The problems encountered by students attempting to obtain performance feedback information from observing Replay with Error Messages were exemplified by documenting a practice approach with a PARTS instructor as controller. Replay with Error Messages was selected after the approach. The replay stopped 18 times to present error messages. In this example, 13 of the 18 error messages (seventy-two percent) were judged to be invalid, 2 were questionable (eleven percent) and 3 valid (sixteen percent). This high rate of invalid error messages is confusing to the student, and it is time consuming. The number of invalid error messages is a multiple problem for training effectiveness. It burdens the student with determining their validity, slows the pace of learning by stopping the replay for each message, and detracts from the instructional value of valid error messages.

Student grades and feedback in PARTS are the product of a sequence of subsystems, beginning with computer speech recognition, and continuing with performance measurement, scoring, grading, and error message selection. The importance of providing accurate performance feedback to a student emphasizes the need for painstaking development of automated subsystems. Any error in speech recognition will be forwarded through the pipeline, and sometimes it will be magnified by the later subsystems. Additionally, the later subsystems may generate their own errors, even if speech recognition is perfect. Therefore, accurate student performance feedback is dependent on the nearly errorless functioning of all the subsystems - a difficult but not impossible task.

The development of an accurate performance measurement system requires refinement of initial models based on empirical data. Initial models of human performance are based on analytical processes such as task analysis and the opinions of subject-matter experts. These techniques provide information for the development of a prototype performance measurement system. At the time of this evaluation, there had been no opportunity for refinement of the prototype PARTS performance measurement system, based on empirical data. Large gains in accuracy would be expected from relatively small improvements in the measurement system.

SECTION IX

USER ACCEPTANCE

User acceptance of new technology is important for its effective application. The phenomenon of "resistance to change" has frequently been experienced with the introduction of new technologies into both operational and training settings. Even a well designed and engineered system can be ineffective if it meets with user rejection and the "not invented here" syndrome (Mecherikoff and Mackie, 1970).

Because they are perceived as relatively exotic, systems that include automated speech technology may be particularly susceptible to the influences of user acceptance. Conversing with a computer is not a routine experience at the present time. Users reactions to AST may be prejudiced by previous exposure to science fiction media, which would tend to generate unrealistically high expectations of the technology. Will such high expectations lead to low user acceptance after the harsh realities of technological limitations are exposed? Will instructors feel threatened or aided by a system that can duplicate some of their functions? These questions were addressed in the present study through questionnaire and interview.

STUDENT EXPECTATIONS QUESTIONNAIRE

The expectations questionnaire was administered to 66 ATC School students after they had completed the ASR course, but before beginning the PAR course. The objective was to survey their expectations about the capabilities of an automated training system, particularly automated speech recognition, before they knew whether they would be selected for PARTS training. Table 17 gives the results of the expectations questionnaire.

The students had high expectations of the automated speech understanding system. Nearly half of them expected the computer to always understand a proper transmission. Those who did not were asked to estimate what percentage of time the computer would understand. Responses ranged from sixty-five percent to ninety-nine percent. The mean estimate of correct computer understanding of voice transmissions was eighty-six percent. It is interesting to note that the transmission recognition accuracy mean from the PARTS P-run records was eighty-five percent.

The need to repeat a transmission for the speech recognition system was expected to occur very rarely or never by twenty-six percent of the students, and occasionally by sixty-five percent. But seventy-nine percent thought they might have to modify their normal speech patterns, such as introducing slight pauses. The likelihood of losing control of the final approach due to repeated speech recognition errors was thought to be unlikely to impossible by fifty-five percent of the students, but another forty-four percent responded "could happen."

TABLE 17. STUDENT EXPECTATIONS QUESTIONNAIRE RESULTS

YOUR NAME N=66 Students DATE ______

Did you see Star Wars? YES 70% NO 30%
How many times? M=2.3*
If YES, do you expect the PAR computer to be as easy to converse with as C3PO? YES 35% NO 65%

When you give a proper transmission, such as "Slightly above glidepath, turn right heading 164, "do you expect that the computer will always understand it? YES 42% NO 58%
If NO, what percentage of the time do you think the computer will understand such a correct transmission? 86%

Do you expect that you will have to change the way you normally speak, such as introducing slight pauses, in order for the computer to understand you? YES 79% NO 21%

How frequently do you think you might have to repeat a transmission because the computer failed to understand it the first time?

NEVER	VERY RARE	OCCASIONALLY	REGULARLY	EVERY TIME	
5	21	65	7	2	%

How likely do you think it would be for the computer to repeatedly fail to recognize a transmission, causing control of the final approach to deteriorate?

IMPOSSIBLE	VERY UNLIKELY	UNLIKELY	COULD HAPPEN	LIKELY	
2	29	14	44	0	%

How do you expect the "voice" of the system will sound?

VERY PLEASANT	PLEASANT	SO-SO	MILDLY IRRITATING	VERY IRRITATING	
0	18	67	14	2	%

*M =MEAN

TABLE 17. STUDENT EXPECTATIONS QUESTIONNAIRE RESULTS (CONTINUED)

How difficult do you expect it will be to understand the computer "voice"?

VERY DIFFICULT	DIFFICULT	OCCASIONALLY DIFFICULT	TAKES SOME GETTING USED TO	VERY EASY	
		2	35	63	%

The PARTS computer may never completely replace an instructor, but it may take over for him in conducting practice sessions and teaching some PAR procedures. How much personal attention and assistance do you think you would get from the instructor? (Check one)

55% 0-5 Times per session
38 5-10 Times per session
7 10+ Times per session

How would you expect to feel about the amount of personal attention you get from an instructor?

VERY BAD	BAD	NEUTRAL	GOOD	VERY GOOD	
0	3	55	31	10	%

How much time would you expect the computer-based training to take (in comparison to normal training methods)?

MUCH LESS	LESS	SAME	LONGER	MUCH LONGER	
5	53	27	15	0	%

At the end of the PAR course, how skilled do you think the average computer-trained student would be (in comparison to students with normal training)?

MUCH BETTER	BETTER	SAME	WORSE	MUCH WORSE	
6	40	45	7	0	%

Check the words that best describe your feelings and attitudes about the automated training program.

38%	Excited	0	Frustrated	17	Skeptical
2	Reluctant	10	Nervous	10	Indifferent
86	Curious	29	Confident	12	Tolerant
52	Enthusiastic	23	Uncertain	2	Uninterested

TABLE 17. STUDENT EXPECTATIONS QUESTIONNAIRE RESULTS (CONTINUED)

Would you volunteer for the computer-based PAR training?

DEFINITELY NOT	NO	MAYBE	PROBABLY	DEFINITELY YES	
0	3	27	23	47	%

How would you feel about being assigned to a group for computer-based PAR training?

VERY BAD	BAD	NEUTRAL	GOOD	VERY GOOD	
0	3	33	36	27	%

Expectations about the sound of the computer generated voice were neutral. Sixteen percent expected it to be mildly or very irritating, but eighteen percent expected it to be pleasant. No students expected that understanding the computer generated voice would be difficult, and sixty-three percent thought it would be very easy.

The students seemed to feel comfortable with automated performance evaluation. None expected to feel bad about it, fifty-one percent were neutral, and forty-nine percent expected to feel good or very good about being evaluated by the computer.

The students' responses did not indicate fears about reduced personal attention from instructors. Ninety-three percent expected to interact with the instructor less than 10 times a day, and ninety-seven percent felt between neutral and very good about the expected amount of instructor attention.

Moderately high expectations of training effectiveness were indicated by the students. Eighty-five percent expected computer-based training to take the same or less time (compared to normal training), and only fifteen percent expected automated training to take longer. Their expectations about the skill level attained through automated training also were moderately high. Forty-six percent expected PARTS-trained students to be better, and only seven percent expected them to be worse than students trained normally. Nearly half (forty-five percent) expected skill levels at the end of training to be about the same, which was the result indicated by the transfer of training study.

An adjective checklist was used to elicit the students' feelings and attitudes about the automated training program. The results revealed positive attitudes mixed with curiosity. The most frequent response was "curious" (eighty-six percent) followed by "enthusiastic" (fifty-two percent), "excited" (thirty-eight percent), and "confident" (thirty percent). The least frequent responses were "frustrated" (zero percent), "uninterested" (two percent), "reluctant" (two percent), "nervous" (ten percent) and "indifferent" (ten percent).

When asked whether they would volunteer for computer-based PAR training, forty-seven percent responded "definitely", twenty-three percent "probably", and twenty-seven percent "maybe." Only three percent responded "no", and nobody responded "definitely not." Similarly, when asked how they would feel about being assigned to computer-based PAR training, sixty-three percent said they would feel "good" or "very good", thirty-three percent responded "neutral" and only "three percent" said they would feel "bad."

In summary, the expectations questionnaire revealed that the students, prior to PAR training, tended to have high expectations of the accuracy of computer speech recognition. They were comfortable with the concept of computer performance evaluation, and they did not expect to resent having less interaction with the instructor. The students tended to expect that automated PAR training would be fast and more effective than traditional methods, and two-thirds reported that they would volunteer for automated training.

STUDENT POST-TRAINING QUESTIONNAIRE

The Post-training Questionnaire was administered to 10 students after they had completed PARTS training, P-run, and the TOT criterion test. The intent of the questionnaire was to provide the students with an opportunity to express their opinions about PARTS training and the associated technologies. The results are given in Table 18. Overall, the students were very positive about their training experience with PARTS. Many "excellent" ratings were given throughout the questionnaire.

Students seemed satisfied with the computer speech recognition, although certain phrases were identified as having recognition difficulty. The Voice Test procedure was used occasionally to often by eighty percent of the students, but Voice Retraining was used much less frequently.

Computer speech synthesis received the highest ratings of any of the PARTS features. All of the students rated the synthezied voice as "easy" to understand. Likewise, all of the students found it easy to identify what speaker was being simulated. One student commented that identifying the speaker (pattern controller, pilot, etc.) was difficult at first, but easy after a little experience.

The instruction received from PARTS generally was rated good to excellent. The material presented on the small CRT was rated slightly higher than the Student Guide, although both were rated very highly. The use of demonstrations of PAR procedures received lower ratings, twenty percent fair and only ten percent excellent. The syllabus was considered good by eighty percent and excellent by twenty percent. The limited amount of direct attention from an instructor was well received by students, with forty percent rating it excellent. One student rated it fair, but additional comments indicated that this case was an apparent personality conflict rather than a reaction to the frequency of student/instructor interaction.

The automated performance measurement system received the lowest ratings in the questionnaire with forty percent of the students rating it fair. The only sub-category of performance measurement and feedback to receive high ratings was the track history (long trails), which was rated excellent by seventy percent. Replay with error messages was selected rarely or never by sixty percent of the students, but this result is heavily influenced by instructors discouraging its use.

The PARTS simulations of PAR display, pilot/aircraft and wind all received very high ratings.

Three summary questions addressed the students' overall rating of their PARTS training. Seventy percent responded that they were glad they received PARTS training rather than the normal classroom and laboratory. The difficulty of transition to the PAR lab (TOT test) was

TABLE 18. PARTS POST TRAINING QUESTIONNAIRE RESULTS (N=10)

COMPUTER SPEECH RECOGNITION

1. Overall rating of the computer speech recognition

POOR	WEAK	FAIR	GOOD	EXCELLENT
		30%	50%	20%

2. What particular phrases seemed to cause recognition problems?

   Turns (Vectors), Range calls, Glidepath position and trend, Handoff Approaching Glidepath, Over Landing Threshold, and Decision Height.

3. Approximately what percentage of time were these phrases not recognized correctly? (Write in a % over the phrases you listed above.)

   Range 5-90%

   Mean = 35%

4. How often was your control of the approach lost because of speech recognition problems?

NEVER	RARELY	OCCASIONALLY	OFTEN	VERY OFTEN
	30%	60%	10%	

5. The training value to you of repeating the PAR phrases in "Voice Training" was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
	10%	20%	50%	20%

6. I used the "Voice Test" procedure:

NEVER	RARELY	OCCASIONALLY	OFTEN	VERY OFTEN
	20%	40%	40%	

7. I used the INIT NEW R/T (Voice Retraining) procedure:

NEVER	RARELY	OCCASIONALLY	OFTEN	VERY OFTEN
	40%	60%		

TABLE 18. PARTS POST TRAINING QUESTIONNAIRE RESULTS (N=10) (CONTINUED)

COMPUTER SPEECH SYNTHESIS (The Voice of Egor)

8. Understanding the computer synthesized voice was:

DIFFICULT	MODERATE	EASY
		100%

9. Knowing who was speaking (Pattern controller, pilot, etc.) was:

DIFFICULT	MODERATE	EASY
		100%

PAR INSTRUCTION

10. The Student Guide was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
			70%	30%

11. The Instructional material presented on the CRT (small screen) was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
			50%	50%

12. The use of DEMONSTRATIONS of PAR procedures was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
		20%	70%	10%

13. The seven-level SYLLABUS was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
			80%	20%

14. The amount of attention from the human INSTRUCTOR was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
		10%	50%	40%

PERFORMANCE MEASUREMENT AND FEEDBACK

15. The automated Performance Measurement System was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
10%	10%	40%	30%	10%

TABLE 18. PARTS POST TRAINING QUESTIONNAIRE RESULTS (N=10)
(CONTINUED)

16. The GRADES of "Safety Error, Needs Work, Satisfactory, and Perfect were:

POOR	WEAK	FAIR	GOOD	EXCELLENT
		40%	50%	10%

17. The TRACK HISTORY (or "Long Trails") of the approach given after each practice problem was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
			30%	70%

18. I used REPLAY with ERROR MESSAGES:

NEVER	RARELY	OCCASIONALLY	COMMONLY	FREQUENTLY
20%	40%	40%		

19. The instructional value of Replay with Error Messages was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
		40%	20%	10%

Comments: Not applicable: 20%

SIMULATIONS

20. The simulated PAR DISPLAY was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
		10%	60%	30%

21. The movement of the simulated PILOT/AIRCRAFT was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
			70%	30%

22. The simulated WIND was:

POOR	WEAK	FAIR	GOOD	EXCELLENT
		10%	50%	40%

TABLE 18. PARTS POST TRAINING QUESTIONNAIRE RESULTS (N=10)
(CONTINUED)

TOTAL TRAINING SYSTEM

23. Are you glad that you received PAR Training on the automated system rather than the normal classroom:

NO	UNDECIDED	YES
	30%	70%

24. Transition to the regular PAR laboratory was:

DIFFICULT	MODERATE	EASY
	60%	40%

25. An overall rating of the PAR training received from the automated training system:

POOR	WEAK	FAIR	GOOD	EXCELLENT
		10%	50%	40%

not rated as difficult by any students; sixty percent rated transition moderate and forty percent easy.

PARTS training overall was rated fair by ten percent of the students, good by fifty percent and excellent by forty percent. This reflects a high degree of user acceptance on the part of the students. A typical final written comment was "Excellent system. With a little more programming it will be far superior to present instructional material and methods."

PARTS STUDENT INTERVIEWS

Twelve students were interviewed after completeing the PARTS syllabus and passing the P-run. The general reaction of most students was positive toward PARTS training, as indicated by the post-training questionnaire. At least one student was more negative in his comments during interviews, however, than when the opinions were expressed more formally in the questionnaire. A summary of the student interviews will be presented by topic.

COMPUTER SPEECH RECOGNITION. There was considerable variability in the students' opinions about computer speech recognition. The most common student opinion was that speech recognition was the main problem with the system. Typical complaints were: speech misrecognitions led to loss of control of the approach, the pilot verbal response early in the approach would frequently give a "Roger" to the previous transmission rather than the present one, or the aircraft would make the turn but not give the "Roger" verbal transmission. One student felt that he had considerable trouble with speech recognition and attributed the problem to his accent. He used voice test regularly and seemed to have trouble with the phrase, "turn right heading." He reported that the aircraft frequently seemed to execute missed approaches when he gave vectors. Another student reported that he had no problem with speech recognition until the P-run. This student was the one, identified earlier, whose P-run came immediately after a software change was implemented that affected voice recognition. Another student reported not having any trouble with the speech recognition system. He commented that he used the voice test and voice retraining functions regularly to ensure good understanding.

VOICE TRAINING. Some students felt that voice training was helpful in allowing them to practice PAR phraseology. Other students felt that it was a necessary drudge; boring, but necessary in order to get on with the automated training. One student commented that voice stylization was not taught very well during voice training, but that it was not too hard to learn.

PARTS INSTRUCTION. A diversity of opinion again was found with regard to student comments on PARTS instruction. One student felt that the CAI information was superfluous. He learned everything from the student guide the night before, and pressed the "Next" button when the instructional information was presented on the CRT. Another student specifically commented that the information on the CRT was very helpful in supplementing the information from the student guide that he had read the previous night.

One student commented that the syllabus was very good in presenting the material in a series of steps rather than presenting it all at once, as in the PAR laboratory. Several other students made similar comments supporting the idea of incremental learning as opposed to the current PAR training method.

The instructional value of the model controller was questioned by one student who said that the model controller "didn't do it right", particularly in omitting "approaching glidepath" calls.

One of the few students who progressed to the enrichment topics in Level 7 commented that servoing was very difficult to understand. He suggested that, if servoing is necessary, it should be taught earlier so the student can practice it at the beginning of each problem.

PERFORMANCE MEASUREMENT AND FEEDBACK. One student commented that the track history was helpful, especially for determining the effects of wind. Most students were favorably impressed by the track history.

Students were somewhat divided on their opinions of the grades given after each practice approach. One student reported that grading was fairly accurate except on clearance, wave-off and decision height where "SAFETY ERRORS" were always given, and range calls where "Need Work" was always given. A frequently mentioned error was that the grade "Perfect" was given when things hadn't been done. Several students gave the example of failing to acquire the handoff, then receiving many "perfect" grades at the end of the approach that they never controlled. Most students reported that they did not trust the grades given by the performance measurement system. Several of them selected replay with errors once or twice in order to determine why certain low grades were given. After observing Replay with Error messages the student concluded that the performance measurement system was error prone, and they tended to discount the grades given at the end of each approach. One student commented that his instructor had told him not to select Replay with Errors but that he hit the button accidently. After observing it once he was glad that he had not used it regularly because of the long time required.

GENERAL COMMENTS. In general, students were quite favorable to PARTS. Nearly all of them reported that they were glad they had worked with PARTS, would volunteer again, and would tell their friends to volunteer. Frequently, students commented that the PARTS training tended to take too long. One said that training time could be shortened by doing five rather than ten practice approaches in each syllabus level. Another student recommended shortening the course to three days.

The most negative comment about the system came from one student who said that working with Egor was "a trip" and he would be glad when he was rid of it. He said that there was no way the computer could teach properly without the instructor there to clarify things. He also felt that he could not trust the performance measurement system. The student said he would not volunteer for this type of a program again because of the problems he had with speech recognition. He did comment favorably about some of the instructional material however, particularly the information displayed on the CRT.

By contrast, the student who worked on the system the same week said that he most definitely would volunteer again. He wished the entire program in air traffic control were like this.

A typical comment by one student was that he enjoyed the experience and looked forward to working with the computer each day. His only negative comment was that voice training was boring. Another typical comment was that the entire experience was great except that voice recognition was sometimes frustrating.

In summary, interviews with students after PARTS training indicated a high degree of acceptance of the system. One student was generally unfavorable toward the system but 11 of the 12 students interviewed had positive attitudes toward their training experience on PARTS.

INSTRUCTORS ATTITUDES AND OPINIONS TOWARD PARTS

The PARTS instructors were an integral part of this evaluation. Their comments and suggestions have been reflected throughout this report. At the end of the evaluation period, final interviews were conducted with the two PARTS instructors, and nine hours of tapes were generated. It is difficult to summarize accurately such a large amount of data, opinion and attitude. The PARTS instructors, in general, seemed moderately impressed with the technological capabilities exhibited in PARTS. They felt that there is definitely a future for these technologies in air traffic control training systems. At the same time, they noted many discrepancies and errors in the prototype PARTS.

The PARTS instructors agreed with the objectives of reducing instructor work load, reducing the instructor to student ratio, and increasing the scope of training through the use of automated technologies including speech recognition and synthesis. Both PARTS instructors commented that the extension of these technologies into other areas of air traffic control beside PAR would be worthwhile, specifically ASR, Radar Air Traffic Control Facility (RATCF) and Carrier Air Traffic Control Center (CATCC).

Although the instructors were impressed with the potential of the technologies demonstrated by PARTS, they had many objections to the details of prototype PARTS training. The instructors would not give their approval to any recommendation that the prototype PARTS be implemented as an operational trainer in its present state. One instructor commented that PARTS currently appears to have so many problems that it might be best to start over again to redesign the courseware. The thrust of his comment was that PARTS has successfully demonstrated the technology but that attempts to patch it would be better served by starting anew, using the lessons learned from the current development and evaluation experience.

Following is a more detailed summary of the PARTS instructors comments from the taped interview sessions.

USE OF COMPUTER SPEECH RECOGNITION IN TRAINING. Both instructors seemed favorably disposed toward this concept. One said that the use of computer speech recognition in training is a good idea, but the entire system must be designed well to make it work. He continued by saying that critical areas for speech recognition must be identified in the design process, such as vectoring for PAR, in order to support the speech recognition process and reduce the impact of errors. The other instructor concurred with the comments and added that reducing training time by nearly 50% by eliminating of unproductive student time spent as "bug operators" is one of the primary advantages to be gained from using computer speech recognition. Both instructors suggested that expanding the use of computer speech recognition to other areas of air traffic control training should be considered.

THE USE OF COMPUTER SPEECH SYNTHESIS IN TRAINING. Both instructors were favorable toward speech synthesis. One commented that there should be different voices for different people involved in the air traffic control task. Additionally, the delay or slowness of the synthesized voice was somewhat of a drawback in the prototype PARTS, and future systems should look toward increasing the speech rate. The other instructor commented that speech synthesis is more personal than CAI, and would be expected to enhance learning by providing a different source of instructional information. He suggested that students do better listening than reading.

AUTOMATED PERFORMANCE MEASUREMENT IN TRAINING. Again, both instructors agree with this concept and emphasize the importance of good design. The need for automated performance measurement to reduce the instructor load was cited by one instructor. Both commented on the importance of the accuracy of automated performance measurement and the need to start with the "conception of the end product", that is, the performance that would be expected from the ideal graduate of the course. In order to achieve this goal, they strongly recommended that a subject matter expert (air traffic controller) be assigned to work with systems designers to achieve an accurate performance measurement system.

THE USE OF AUTOMATED SYLLABUS CONTROL. One PARTS instructor advocated what he called semi-automated syllabus control, or flexibility in syllabus control. Accurate information about student performance should be available to the instructor so he can assist in syllabus decisions, and exercise an override function if necessary. These comments refer to the more global aspects of syllabus control such as advancing to a new level of the syllabus or remediating the students for review. Some of the finer details of problem generation were not discussed in this interview. But it is obvious that instructors with multiple student stations would have an impossible task attempting to initialize the parameters for each practice run. Automated syllabus control with the capability for instructor override seems to be the plan advocated by the instructors. One instructor suggested that an ideal design would be to automatically adjust problem difficulty when a student is having trouble, then provide automated remediation if necessary, and, finally, alert the instructor if the student continues to perform poorly. Instructor overrides would be available throughout this adaptive remediation process.

INSTRUCTOR ROLE IN AUTOMATED TRAINING SYSTEMS. This issue is summarized by one instructor's comment that, "the instructors role in automated training systems has never been defined." He advocates early definition of the optimum instructor role so that system design can complement the instructor functions. Feedback to the instructor about student performance is critically important in order to facilitate any role of the instructor in working with an automated training system. The instructors feel that their role should be in the area of providing added information that is not available from the automated instruction, such as answering questions for individual students. The instructors have doubts about when it is proper to override the system functions and when to let the system and student work without intervention. Both instructors feel that the use of automated training can increase the effectiveness of instruction in air traffic control, including using more variables such as different aircraft speeds and pilot skill levels.

In addition to these extra variables, automated training can reduce the instructor workload in the area of moment by moment performance measurement and repetitive instruction on basic techniques. However, the integration of automated systems with new instructor roles remains a question.

There is a great need for "PR", as one instructor said. Other instructors in the air traffic control community should be exposed to the possibilities of automated training for easing their workload while changing their role. This suggestion for the need of public relations in the area of introducing automated training concepts is similar to the conclusions reached in a study of the acceptance of technological innovations in the Navy (Mecherikoff and Mackie, 1970).

PARTS TRAINING EFFECTIVENESS ISSUES. The interviews with PARTS instructors turned from general issues of automated training systems toward more specific issues dealing with PARTS.

A summary of the attitudes of the PARTS instructors toward PARTS training effectiveness can be given succinctly. One instructor stated that PARTS, "as it is now, would not work as an operational system." The other instructor summarized by saying "over all, the system has lots of problems." The examples cited as evidence for the overall opinion included the following: if the prototype PARTS were an operational system, the students would do Replay with Errors as well as Freeze and Feedback, and never finish the course; there were many errors in the courseware; the vocabulary showed that the instructional material was not designed by air traffic control people; ties between the student guide and the CAI instruction on the system were not smooth; the training emphasis was on handoffs rather than on conduct of an actual approach; some of the demonstrations used in conjunction with voice training were incorrect; and the model controller was a great idea for demonstrating concepts, but "it" did not work.

Very brief summaries will be given of the instructors comments regarding the training features of PARTS. Detailed discussions of these features are not possible in the present report because of the length of comments provided by the instructors.

Instruction. The tutorial aspects of PARTS received mixed reviews. The student guide was considered to vary from adequate to very good. The background information and the level of detail in describing procedures were good. Certain instructional areas within the student guide were considered weak, such as approaching glidepath and begin descent. The instructors felt that more input to the student guide was needed from an expert in air traffic control, to correct some of the errors in content terminology.

The CAI type of instruction was considered acceptable in most cases. However, some of the material was described as "awfully confusing." Students sometimes did not know when to push a button, or which button to push. Again, some of the content areas were weak in the CAI, such as approaching glidepath.

The layered instructional approach was appreciated by the instructors. The gradual introduction of new PAR concepts gave students a chance to practice and assimilate subsets of the total PAR task.

Model Controller. The model controller, and his demonstrations, were considered a good idea but "less than model, for sure." The instructors felt that the concept of a model controller was excellent to provide the student with demonstrations of how the procedures should be done. The concensus opinion on the model controller was that it is a great idea but needs a lot more work. "He failed the P-run."

Syllabus and Syllabus Control. As mentioned, the syllabus was judged to be a good concept in the sequential presentation of materials. However, certain procedures such as glidepath position and trend should occur sooner in the syllabus. That procedure is currently not introduced until the third day, resulting in too few practice approaches with both glidepath and azimuth. The syllabus has some problem with uneven level of detail. A great deal of time is spent on vectoring, but the student already has been exposed to vectoring for three weeks in ASR.

Syllabus control was thought to be a good idea but it was not properly implemented. The instructors felt that they had to override constantly. In situations where no override was possible they had to tell the student "don't look at this." One instructor commented that in a perfectly designed system, instructor override functions would not be necessary. In a multiple station version of PARTS it would be impossible for instructors to go around to each student telling him when not to believe the instructional material. Likewise, instructors could not exercise override functions regularly. Syllabus control would be a very important aspect of a multiple student station operational version of PARTS. The instructors felt that considerable work was necessary to enhance the PARTS adaptive logic and performance measurement systems to provide for individual learning needs.

Computer Speech Recognition in Prototype PARTS. The two instructors emphasized different aspects of computer speech recognition in PARTS. One emphasized the use of computer speech recognition as the "thing of the future" in training systems. He thought some benefit resulted from forcing the student to articulate properly. User acceptance of speech recognition was seen as the key issue. Students were sometimes frustrated by approach control problems when speech recognition errors occurred.

The other instructor stated that generally, the speech recognition in the prototype PARTS would require improvement for a successful operational system. He emphasized the differential recognition accuracy needed for various transmissions in the PAR task. For some transmissions, ninety percent recognition accuracy, or less, may be acceptable. But for other transmissions very high recognition is necessary, possibly ninety-nine percent. The instructor also pointed out a subtle training problem induced by speech recognition errors in vectoring. If a student gives a turn and it is not recognized, he may give the same turn again when he suspects that the speech recognition system did not recognize his first transmission. Giving two consecutive vectors to the same heading is not acceptable procedure. The PARTS instructor felt that students were attempting to accommodate PARTS recognition errors and, consequently, exercising poor PAR technique. In summary, this PARTS insructor felt that further development of the speech recognition software is necessary to prevent important misrecognition errors. He added that the recognition problems in the prototype PARTS do not seem to stem from the IWR technology, but from "fixable" software that applies the technology to the PAR task.

The voice training required for speech recognition was mildly criticized by the instructors. They agreed that a few repeats of each phrase are helpful to students for learning PAR phraseology, but ten repeats of a phrase lengthens the procedure to the point of being boring to the student and wasting training time. Voice training in the context of approach control is good, and was well done in the prototype PARTS, but more context usage of phraseology would be desirable. One instructor suggested the possibility that the voice test function also should be done in context of the approach.

The speech recognition system seems to work well for experienced users. It takes time to learn speech stylization. By the time most students are stylizing correctly and getting good recognition accuracy, the PAR training course is nearly over. One instructor suggested that more time should be spent instructing the student how to use the speech recognition system. He noted that the prototype PARTS curriculum included little or no instruction on stylization. Nearly all the emphasis in both voice training and voice test was on each phrase individually. He suggested more intensive training early in the course on speech stylization techniques.

Performance Measurement System. The automated performance measurement system was described as "having problems." One instructor described PMS as a study in itself. The instructors suggested that the entire performance measurement system needs to be reevaluated. The weighting factors for performance categories should be reviewed, the subtractive technique of scoring should be reconsidered, and techniques for reducing the impact of speech recognition errors on performance measurements should be studied. The performance measurement categories are not differentiated, for the student, but some performance categories are much more important than others. The addition of the "SAFETY ERROR" grade during the evaluation period was a step in the right direction.

Another aspect of performance measurement that needs further study is the nature of the information presented to the instructor. Giving the instructor a number representing the student's performance in some performance category may or may not be sufficient. A numeric value must be meaningful to the instructor in order for him to understand the student's progress by examining automated records. Alternative techniques could be used, such as simply listing the frequency of errors made in each category. This evaluation did not clarify the types of instructor information that would be most useful for a situation involving multiple students. The one-to-one student-to-instructor ratio in the present study gave substantial time to the instructor to interpret performance measurement scores of the student. This time would be severely reduced in a multiple-student PARTS.

One PARTS instructor emphasized the relationship between instructor information requirements and the adequacy of the automated performance measurement and syllabus control The instructor's need for information about the student is greatest when the automated adaptive instruction is <u>not</u> working properly. In the perfectly designed system the instructors information requirements would be low because he would rarely be required to make decisions about syllabus advancement or remediation.

<u>Performance Feedback</u>. Performance feedback to the student elicited comments from the instructors that were very similar to the student comments discussed previously. The instructors felt that the track history (long trails) were excellent. The grades presented to the student after each practice problem were described as "misleading." The grades were thought to be too simplified. The students were capable of understanding more detailed information, such as point totals or bar graphs, as suggested by one instructor.

It is difficult to separate the instructor's criticisms of the verbal grades given to the students from the fact that the accuracy of the performance measurement system led to incorrect grades being given. Their criticism of the verbal grades, given by performance category, may in large part be directed at the inaccuracy of the grading system itself. Clearly, no type of feedback presentation to the student will be satisfactory if the performance measurement system is not accurate.

<u>Replay</u>. Replay and replay-with-error-messages, as sources of feedback to the student, were severely criticized by the instructors. Replay alone for the PAR task seems unproductive, according to one instructor. He suggested, however, that a replay that included recognized phrases might be helpful to the student. One instructor described replay with error messages as potentially an excellent teaching tool, but "it doesn't work right." The other PARTS instructor described replay with

errors as "awful...there is nothing good about it." He described replay with errors as "a fantastic idea" if it were not so time consuming, and if the errors made sense. Another suggestion for improvement in the replay with errors function was to enable the student to take a "quick look"; to observe a replay with error messages only for one portion of the approach in which a particular error was made. The other instructor concurred, suggesting that a selectable starting point for replay with errors would be ideal. In the same vein, he suggested that error messages should be selectable for any performance category, rather than having to watch all of them. Further suggestions included making the error feedback more interactive. For example, an error message could be given to the student with a menu of possible student actions, such as to obtain a simple explanation on the CRT, to read certain sections of the student guide, or to call the instructor. In other words, rather than merely listing errors for a student, the system should give him guidance about remedial measures to avoid making the same error in the future.

Aircraft/Pilot/Environmental Model. The aircraft/pilot/evironmental model was described by the PARTS instructors as basically good, although a considerable number of problem areas were identified. Having four types of aircraft is excellent training for students. The aircraft responsiveness seems good, except that turns and begin-descent are too abrupt.

The five ability levels of pilots were thought to be more complicated than necessary, by the PARTS instructors. Neither students nor instructors could identify the bad pilot. Therefore, the instructors felt that the student received little or no training value from the variable skill levels of the pilots. One instructor suggested that the system should tell the student when a bad pilot is going to occur, as an occasional instructional device. One final comment about the pilot model is that the pilot is too good on glidepath, even the Number 4 (next to the worst) pilot. This instructor felt that PARTS students generally were weak on glidepath procedures because the pilot model was too good at maintaining glidepath. One instructor suggested that the pilot/aircraft occasionally should be shown arriving on the PAR display at less than nine miles.

The wind model was thought to be good by both instructors. The inclusion of wind variables such as gust and anti-gust seemed overly sophisticated for basic PAR training. When the wind model was working correctly, it was excellent for training. However, sometimes it was not working properly and the aircraft seemed to be deviating from course in a direction toward the wind. The instructors suggested that other variables which affect the PAR task may be more valuable to simulate than the elaborate pilot and wind models. Examples of these other variables are density altitude, inversions, ground clutter, and more than one aircraft on the scope simultaneously.

Adaptive Problem Difficulty. The adaptive assignment of problem difficulty by manipulating aircraft type, pilot skill level and wind was thought to be a good idea but not well designed. The slowest aircraft, the U-21, was overly used throughout PARTS, while the P-3 aircraft was never encountered in the Levels 4, 5 and 6 of the syllabus. One instructor reviewed the computer records of aircraft assignment and commented that the student "sees that Army aircraft all the time." Because the Army U-21 was so slow, the instructors attempted to override this aircraft as much as possible in syllabus Levels 5 and 6. One instructor commented that the wind variable was the only helpful item in the adaptive process. Students having difficulty with vectors during cross wind were helped by reducing the cross wind factor during practice. In summary, both instructors felt that a well designed adaptive process for assigning practice problem difficulty variables would be a good training enhancement.

Remediation. On the topic of remediation, one instructor commented that "there was none." Both instructors felt that a remediation process for students having difficulty would be an excellent idea. It could be automatic syllabus branching, or it could be an instructor option, given that sufficient diagnostic information was available to the instructor. A freeze and feedback mode was suggested as a good candidate for a remedial assignment. Additionally, remediation could take the form of studies in the student guide, special handouts, or consultation with the instructor.

User Acceptance. User acceptance was described by one instructor as critically important. He described a background of discomfort and possibly resentment toward the prototype PARTS from other instructors at ATC Schools. He felt that they were not properly briefed on the system and therefore its arrival on the scene was interpreted as an invasion on the part of instructors not directly involved with it. This lack of acceptance on the part of other ATC School instructors was not enhanced by the early system breakdown problems. The prototype PARTS quickly developed a reputation for being unreliable.

Both instructors described student acceptance as fairly high. Students reported to the instructors that they liked the system and enjoyed the experience of working with it. One instructor described his personal acceptance of the current version of PARTS as "very low." The relatively low acceptance on the part of the PARTS instructors was described as developing from the need to override and interfere continually with the system in order to make it work. Rather than reducing instructor workload, it required full time work on the part of the instructor in order to get one student through the system in the time allotted. One instructor commented that the system was designed to make PAR training nearly instructorless, but he has never worked so hard. The other instructor emphasized the need for user acceptance in the broadest sense, including all instructors at the Air Traffic Control Schools, as well as students. He repeated the need for a program to introduce the system, sell the system, communicate with all the users and potential users, and generally do a good job of public relations.

Device 15G19 Retrofit. Discussions regarding the issue of retrofitting the 15G19 demonstrated that the instructors were able to separate their opinions of the technologies represented in the prototype PARTS from their discouragement with some of the specific features in it. Both instructors supported the concept of introducing the automated technologies of PARTS into other areas of air traffic control training.

In ASR and RATCFF students spend at least half of their time operating the TCC ("flying the bug") for the benefit of the controllers training. The PARTS instructors stated that the primary benefit to be gained from automating ATC training would be to eliminate the time spent by students in this training support role. This time savings could be translated into reduced training time or more intensive controller training in the time currently allotted.

Retrofitting the 15G19 should include all aspects of the 15G19, not just PAR. One instructor suggested the possibility of a hybrid system which would replace the bug operator with speech recognition and modeling the pilot/aircraft, while retaining instructors in their current role of providing instruction and performance measurement. Eliminating the time spent as bug operator was seen as the primary rationale for the implementing technologies demonstrated in the prototype PARTS. Upon further reflection, the PARTS instructor recommended against the hybrid concept because the current instructor cadre would be faced with twice as many students in the PAR course. The hybrid system would eliminate inefficient use of time as a bug operator, but it would increase instructor workload by one-hundred percent.

Shortening the five day PAR training was not seen as a constructive objective by the PARTS instructors. One reason that the training time could not be significantly reduced in PAR, according to one instructor, was that the time required for voice training somewhat replaces the time required for bug operation. Applying the automated technologies to other aspects of air traffic control, such as RATCF and ASR would allow the automated technologies to reduce training time by eliminating the requirement for bug operators. The other PARTS instructor wondered if a 15G19 retrofit could include the use of the present PAR display consoles, rather than using the console of the prototype PARTS. A stand-alone multiple station PARTS, rather than a retrofit, would be aceptable, said one instructor, if everything worked right. A multiple station version of the current prototype PARTS was suggested by the interviewer, and the replay was, "No way. Absolutely not." Improvements in the courseware and the software supporting speech recognition would be essential for an operational multiple station PARTS.

INSTRUCTORS' CONCLUSIONS. The instructors concluded their overall evaluation of the prototype PARTS with the following comments: "The technologies are here." "The entire system needs to be redesigned and we must continue to get feedback from subject-matter experts for changing the system." "More data needs to be collected about the role of the instructor and the transfer of training of PARTS to actual application in the fleet." "The design should be flexible for easy modification as ideas are tried." "Subject matter experts should be available continuously during the system development process." "The automated training system must have the capability for the instructor to override it at any point." "The automated technologies are good, and with modification and redesign, we could have a system that would be very helpful."

INSTRUCTOR QUALIFICATIONS AND TRAINING

Implementing automated technologies in training is likely to have some impact on the instructor qualifications necessary for proper use of the system. Similarly, the training requirements for new instructors will be changed. The present evaluation provided very little information regarding these issues. A perfectly operating automated system could have the effect of reducing the expertise required of an instructor. But the experience in the PARTS evaluation indicated the contrary, that besides full expertise in PAR, the instructors needed additional expertise in systems management and operations.

Automated training will not merely reduce the instructor workload, if properly designed, but will change the nature of the instructor task. The instructor is envisioned as a learning supervisor, involved in the higher level of conceptual achievement of the students, rather than the more mundane and repetitious aspects of learning the task. In addition to acting as learning supervisor, he must know the capabilities and limitations of the automated system. He must know when to intervene in the learning process in order to optimize instruction. The student should not be given sole responsibility for initiating student/instructor contact. In order for the instructor to know when he is needed for person-to-person contact, he must have access to accurate summary information about the student's progress. These issues are discussed more fully in a study by Joplin (1980).

Based on the experience gained in the present evaluation, instructor qualification requirements cannot be reduced for an automated system. On the contrary, instructor qualification requirements may have to be more stringent in order to select qualified PAR instructors who also have the capability and desire to learn systems management and operation.

Instructor training for automated training systems is another issue that will need to be resolved. Instructors at Air Traffic Control Schools currently undergo approximately three weeks of training on general instruction and a brief period of apprenticeship before becoming qualified to instruct. This period of instruction would undoubtedly be lengthened for an automated system because the instructor must learn his new role as learning supervisor, as well as becoming familiar with the capabilities of the automated system, the procedures for working with the computer, decision strategies for overriding system functions, and the availability of various tutorial and remedial techniques.

In summary, the evidence obtained from this evaluation, although limited, indicated that instructor qualifications will be the same or higher, and the instructor training procedures will be lengthened for automated training systems such as PARTS.

SECTION X

TRAINING EFFICIENCY

INTRODUCTION

Training effectiveness and user acceptance are central issues in designing and evaluating new training equipment and methods, and integrating them into training programs. The efficiency of new training equipment and methods also can be a central issue, particularly if training resources are constrained or it is necessary to select the most efficient of several equally effective alternatives.

Training efficiency refers to the rate at which resources are consumed in fulfilling training objectives. Resources include development, acquisition, operation and support, and personnel resources. Frequently, costs associated with these resource categories can be established so that cost-benefit tradeoffs can be made among competing training systems. In other instances, cost may be a secondary consideration, as when personnel resources are very limited. Table 19, adapted from Allbee and Semple (1980) lists a comprehensive set of cost categories associated with the acquisition, use and maintenance of training devices.

OBJECTIVE

Training efficiency is an important issue in the present study because of a growing shortage of skilled PAR controllers who can be relieved of fleet duty to serve as instructors at the Air Traffic Control Schools. Thus, primary and practical measures of the efficiency of training using PARTS were:

1) The extent to which the PAR course instructor cadre might be reduced through automated instruction;

2) The extent to which PAR training productivity might be enhanced, primarily by relieving students of the need to function unproductively as "bug operators"; and

3) Estimates of the extent of instructor cadre reductions and improved training productivity that might be achieved in other, similar training programs by applying the instructional technologies and features demonstrated in PARTS.

METHOD

A comprehensive training efficiency analysis of PAR training comparing the use of PARTS technology with present training practices and equipment was not intended as a part of this study. Such an analysis would have required collecting and analyzing detailed cost data in many if not all of the cost categories listed in Table 19. Obtaining meaningful cost data for the numerous cost elements in the various categories is a difficult and often subjective process. This is true even when historical cost data are used, because military cost accounting systems historically have not designed to collect training cost data at the levels of detail needed for highly detailed cost benefit analyses (Allbee and Semple, 1980). Investigations at Headquarters, NATTC confirmed that needed cost data for a detailed analysis of the present training system were not available in the required form and detail from the Navy accounting system. Thus, considerable investigation and estimation procedures would have been required. In addition, PARTS is an experimental prototype device. Acquisition costs for a production run of similar systems would, at best, have been rough approximations. In addition, operation and support costs would have to be estimated based upon Navy experience with similar computer-based systems. This, too, would have required investigations that were beyond the scope of the present study.

The approach taken, therefore, was to focus on efficiencies related to personnel utilization, i.e. reductions in the instructor cadre and more productive use of student time. The first step was to determine whether replacing present training devices and practices with automated PARTS-type technology would allow a reduction in the PAR instructor cadre, and if so, how much of a reduction. This had to be an estimate because the training effectiveness experiment did not allow empirical determination of minimum practical student to instructor ratios. The estimate was made in concert with the instructors who had worked with PARTS before and during the training effectiveness evaluation and who previously had been instructors in the PAR lab. It was their consensus that the normal cadre of four instructors could be reduced to two, for a savings of two instructor positions, or fifty percent. This reduction assumes that many of the refinements to PARTS discussed previously in this report would be implemented. The instructors also pointed out that further reductions in PAR instructor staffing would not be possible due to the numerous administrative tasks that instructors must perform in addition to teaching. Since this is an untested assumption, only moderate confidence should be associated with it.

A majority of PAR instructors hold an E-6 rating. The Composite Standard Military Rate Table (1 October 1979) was used to determine annual billet costs (wages and overhead) for the instructor billets. Instructors at the Air Traffic Control Schools estimated that approximately 3.5 weeks of the total instructor training program is devoted to how to teach the PAR course. Financial Analysts at Headquarters, CNTT (CNTT-N-212), determined that this 3.5 week period of instructor training cost the Navy $1,805. This figure together with billet costs was used to compute cost savings resulting from possible instructor cadre reductions. Additional instruction training costs associated with learning to effectively employ PARTS-like technology

TABLE 19. TRAINING PROGRAM COST CATEGORIES AND ELEMENTS

ACADEMIC COST ELEMENTS

Classroom space Supplies (desks, tables, etc.) Civil Engineering support (prorated per sq ft of space) Expendable training materials (workbooks) Classroom training aids	Usually very low unless new facilities required

Instructor hourly costs (including overhead costs)
Instructor costs
Student wages

TRAINING DEVICE COSTS - Major Components

Acquisition cost
Logistics/depot support cost
Operation & maintenance cost

ACQUISITION COST ELEMENTS

Initial investment
Government procurement costs
Test & evaluation costs

LOGISTICS/DEPOT SUPPORT COST

Initial spares
Spares Replenishment
Depot maintenance
Class IV and V modifications
Contract engineering
Support group costs

OPERATING AND MAINTENANCE COST

- Maintenance
- Operations
  - Program & syllabus development & management
  - Instructor hourly costs (including overhead costs)
  - Instructor training costs
- Civil Engineering support costs
  - Utilities/energy
  - Facilities maintenance & modification
  - Janitorial services
  - Civil Engineering management & engineering support
- Facilities (building) costs
- Student wages
- Instructor costs

were not estimated or included in this computation.

The training efficiency information that follows presents cost savings that could result from reduced instructor cadre size and more productive use of student time. Resulting potential cost savings are presented so that planners can have an approximation of the likely savings from personnel factors alone. Total life cycle cost savings could be greater if all relevant cost categories, such as maintenance and replenishment spares, were taken into account. Therefore, the present method of analysis may have resulted in conservative cost saving estimates. Also, cost savings may be a secondary consideration in relation to other driving factors, such as the need for trained personnel in fleet operations.

INSTRUCTOR CADRE REDUCTIONS

Members of the PAR instructional staff estimated that, with an improved version of PARTS, the instructor cadre could be reduced by two positions (fifty percent). As mentioned previously, this savings should be interpreted with moderate confidence. The typical PAR instructor holds an E-6 rating. The annual billet cost for a Navy E-6 is $15,001.

Historically, instructors have served four year tours at the Air Traffic Control Schools. Headquarters, NATTC estimated that instructor training for the PAR portion of the course was $1,805. Thus, the average cost of instructor training on an annual basis is $1,805 ÷ 4 = $451 per instructor. The average annual cost of an instructor, therefore, is: $15,001 + $451 = $15,452. Annual savings resulting from an instructor cadre reduction of two positions, therefore, would be $30,904. Assuming a 20 year life for PARTS-like training devices, and using the commonly accepted seven percent per year discount factor (i.e. cost escalation factor), savings over a 20 year life cycle resulting solely from instructor billet and training costs would be $1,266,926. Obviously, the use of different discount factors will produce different potential cost savings. For example, if a ten percent discount factor is used, the potential savings over a 20 year life cycle would be $1,770,025.

STUDENT PRODUCTIVITY IMPROVEMENTS

In the present five day PAR course, one day is spent in classroom instruction, and the remaining four are spent in the PAR laboratory. Of these four days, each student spends fifty percent of his time in hands-on training at a PAR scope and fifty percent acting as a "bug operator" for fellow students. Thus, two of the four days are spent performing a function that has little direct relationship to learning to be a PAR controller. The PARTS instructors considered the time spent as "bug operator" to be unproductive with respect to PAR controller training.

A potential saving can be achieved through improved productivity by eliminating the need for "bug operator" tasks. On the average, each PAR class consists of 14 students, each of whom spend the equivalent of two days performing "bug operator" tasks. A new class starts each week. Thus, 28 days per week (per class) of student time is spent performing a task that has little direct relationship to learning to be a PAR controller. If this time could be spent on more future task training, 1,456 mandays per year would be salvaged. Dividing 1,456 by a 365 day year results in a 3.99 manyear per year potential savings simply due to the automation of the "bug operator" function in this one training setting. Using annual billet costs for the E-2 rating of $8,476, the savings in one year alone could amount to $33,819. Again assuming a 20 year training device life cycle and a seven percent discount factor, potential productivity savings for student time are $1,386,378. Assuming a 10% discount factor, the potential life cycle student billet cost savings are $1,937,152.

COURSE LENGTH REDUCTIONS

A few years ago, the PAR training course was reduced in length from three weeks (15 working days) to one week (5 working days). Further significant reductions of this magnitude do not appear warranted. Members of the PAR instructional staff estimated, however, that the present five day PAR course could be reduced to a four day course if PARTS-type devices were used and if the necessary system design improvements were implemented.

Assuming an average of 14 students per week, and a one day per student saving, a total of 728 student training days per year could be saved. This computes to 1.99 student years saved per year. This, in turn, computes to a saving of $16,104 per year, due solely to a reduction of billet costs for students while in training. Assuming a seven percent discount factor, potential productivity savings for student time saved over a 20 year system life cycle are $660,102. Assuming a ten percent discount factor, the potential life cycle savings are $922,437.

Potential savings due to reductions in student attrition were not considered in this or prior computations because student attrition is very low (around 1%) during PAR training. Attrition typically occurs during prior phases of hands-on training in the Air Traffic Control Schools.

PARTS TOTAL POTENTIAL SAVINGS

The preceding training efficiency analyses were limited to potential savings associated only with personnel utilization, as previously discussed. Within this focus, the following conclusions can be drawn, and similar conclusions can be made for other applications of computer speech technology and instructional support technolgies incorporated into PARTS-like devices.

1. Instructor resource savings can be achieved, and the annual (1980) savings for E-6 instructors are $15,452. Assuming a ten percent annual discount factor for 20 years, potential savings are $885,090. These amounts are doubled if, as predicted with moderate confidence, two PAR instructor billets could be deleted.

2. Student productivity savings come in two forms: a) better use of student time; and b) reduced course length. Other potential savings exist if student course attrition rates can be reduced; this factor was not addressed in this study because the attrition rate during PAR training is virtually zero. Assuming no attrition rate differences due to PARTS technology introductions, it is estimated that better use of student time could result in a 20 year life cycle savings of $1,386,378 using a seven percent discount factor, and $1,937,152 using a ten percent discount factor. Reduction of course length by one day could lead to 20 year life cycle cost savings of $660,102 using a seven percent discount factor. Again, these savings result only from reductions in student billet costs, and no other factor.

3. Combining the potential savings identified in 1 and 2 above, a total 20 year life cycle savings of $3,313,406 appears likely using a seven percent discount factor. The potential savings grows to $4,629,614 if a ten percent discount factor is used. These figures are for labor saving only.

OTHER SCHOOL POTENTIAL SAVINGS

This analysis has focused on PAR training. However, discussions with members of the instructional staff lead to the conclusion that the use of voice technology could have significant cost and other resource impacts on other training courses at the Air Traffic Control Schools. For example, Carrier Air Traffic Control Center (CATCC) training involves a five week training course in which one instructor is assigned for each student in laboratory practice sessions. Potential savings may be possible through automated instruction, to the point where several students could be assigned to one instructor. Again, this reduction potential must be assigned only a moderate confidence but since it remains to be demonstrated empirically, However, savings due to reductions in CATCC instructor cadres can be computed as described above, and could be significant over the training system's life cycle.

Along similar lines, the Radar Air Traffic Control Facility (RATCF) course is six weeks in length. Five of these weeks are spent in the traffic control laboratory. During training, a cadre of six instructors takes a class of 14 E-2 students all the way through the course. During the course, approximately seventy-five percent of each student's lab time is spent as a "bug operator" for fellow students. Thus, 52.5 student-weeks per course are spent as "bug operators" (14 x 5 x .75 = 52.5). With a course length of six weeks, 8.7 courses per year are taught. Thus, on an annual basis, 457 student-weeks or 8.7 student-manyears of labor are spent performing a task that has little direct relationship to the achieving of course performance objectives. Using the student billet cost data from prior analyses, a 20 year life cycle cost of $3,022,644 using a seven percent discount factor could be saved by applying PARTS-like technology. Using a ten percent discount factor, the potential savings in student billet costs, alone, rises to $4,223,884.

TRAINING EFFICIENCY CONCLUSIONS

It is quite probable that considerable savings may result from applying modern automated instructional support technology. Additionally, well designed automated instructional devices hold the potential for freeing critical (instructor-level) skills for fleet use. This evidence suggests that automated instructional support technology merits further development and application.

SECTION XI

CONCLUSIONS AND RECOMMENDATIONS

It is difficult to draw conclusions about the application of automated speech technology to training systems from evidence pertaining to only one specific prototype system, PARTS. Given this caveat, the following conclusions and recommendations are based on the available information including subjective opinion of students and instructors, observation of the technology in an applied training environment, cost estimates based on personnel utilization, and a limited amount of experimental/quantitative data.

CONCLUSIONS

COMPUTER SPEECH RECOGNITION IN TRAINING. Computer speech recognition is sufficiently advanced to begin applying it in appropriate (speech related) training tasks. Isolated word recognition has some limitations, namely, the requirements for speech stylization and extensive speech sampling. These limitations can be minimized by careful courseware design, emphasizing stylization instruction and speech sampling within a task-oriented instructional context.

Automated speech recognition enables simulation of a control system that includes verbal advisories or commands, such as air traffic control. Furthermore, it expands the applications of CAI by allowing voice interaction with instructional systems. Another set of automated training functions, not necessarily related to speech recognition systems, includes adaptive or fixed syllabus control, problem generation, record keeping, and performance feedback to students and instructors. Combining the capabilities of automated speech recognition with other automated training functions enables the development of fully automated training systems for speech related tasks.

A summary of the state of automated speech recognition is beyond the scope of the present study, but reviews are available in Dixon and Martin (1979) and Lea (1980). Full detail regarding the PARTS speech recognition system can be obtained from the referenced technical reports on the GCA-CTS development.

From a training effectiveness perspective, the major problem encountered in PARTS speech recognition was the occurrence of a critical recognition error that sometimes caused loss of control of the approach. This problem was indicative of an evolving prototype system rather than an inherent limitation of the technology. Revision of the software supporting PARTS speech recognition would eliminate the problem of critical recognition errors. This issue serves to emphasize the importance of developing task-specific speech recognition software for each application of automated speech technology.

The potential benefits to be derived from well-designed automated training systems with speech recognition include cost savings, the reduction of instructor workload, reduction of instructor to student ratios, elimination of training support personnel who are the recipients of the verbal information, increased student interest, and enhanced training effectiveness through individualized instruction and the systematic manipulation of a wide range of task variables by modeling and simulation.

COMPUTER SPEECH SYNTHESIS. Synthesized speech is easily understood. It can be applied in training systems to capitalize on students natural language skills, without emphasizing reading, and to simulate the verbal communication which may be a source of information necessary to perform a task.

TRAINING FEATURES OF THE PROTOTYPE PARTS. As a prototype training system, the PARTS successfully demonstrated the potential for integrating automated speech recognition, speech synthesis, performance measurement, syllabus control, record keeping, task simulation, and interactive instruction. No deficiencies inherent in these automated technologies were found that would preclude their use in an operational training system. A considerable number of deficiencies were found in courseware design. Although the basic technologies seem adequate, major revisions would be required in the design of courseware including instructional concepts, performance measurement and feedback, and syllabus control, in order to achieve training effectiveness with reduced instructor workload.

TRAINING EFFICIENCY. Well designed training systems utilizing the technologies demonstrated in the PARTS could enable savings in terms of student time, course length reductions, and instructor cadre reductions. The total savings for a five day PAR-type of course are estimated at 4.6 million dollars over a 20 year life cycle using a ten percent discount factor.

RECOMMENDATIONS

1. The feasibility of applying automated speech technology to training systems on a wider basis at the Air Traffic Control Schools, including ASR, RATCF, and CATCC as well as PAR, should be studied (see pages 19, 101-102, 108, 116-117).

2. The instructor's role in automated training systems should be defined during the functional design stage of system development (see pages 55, 103).

3. A subject-matter expert with experience as an instructor should be intimately involved throughout the development of automated speech technology training systems (see Section IX).

4. Research is needed to define the optimal allocation of training control to student, instructor and automated system (see pages 69-74).

5. In the development of AST training systems, emphasis should be placed on the careful integration of the speech technology with other subsystems to prevent the cascading effect of errors in speech recognition (see pages 15, 87).

6. Improvements in the application of computer speech recognition technology to training systems should be sought in the reduction of time required for speech sampling (voice training) and in the reduction of the probability of costly (task-critical) recognition errors (see pages 15-16, 30-44, 99, 105).

7. The accuracy of automated performance measurement is critical to the proper functioning of an automated training system because it becomes the source for student feedback and for adaptive syllabus control. We recommend careful development of automated performance measurement systems through an iterative process of empirical validation and modification (see Section VII and pages 15, 83-87, 106-107).

REFERENCES

Allbee, K.E. and Semple, C.A. Aircrew Training Device Life Cycle Cost and Worth of Ownership. Technical Report AFHRL-TR-80-37. Air Force Human Resources Laboratory, Wright-Patterson AFB, OH, 1980.

Barber, G.C., Hicklin, M., Meyn, C., Porter, J.E., and Slemon, G. K. Ground Controlled Approach Controller Training System. Technical Report NAVTRAEQUIPCEN 77-C-0162-2. Naval Training Equipment Center, Orlando, FL. April, 1979.

Breaux, R. Training Characteristics of the Automated Adaptive Ground Controlled Approach Radar Controller Training System (GCA-CTS). Technical Note NAVTRAEQUIPCEN TN-52. Naval Training Equipment Center, Orlando, FL. July, 1976.

Breaux, R. and Goldstein, I. "Developments of Machine Speech Understanding for Automated Instructional Systems." In Eighth NTEC/Industry Conference Proceedings, New Concepts for Training Systems. Naval Training Equipment Center, Orlando, FL. November, 1975.

Broadbent, D.E. "Levels, Hierarchies, and the Locus of Control" Quarterly Journal of Experimental Psychology, 1977, Volume 29, 181-201.

CNTECHTRA. Curriculum Outline for Air Controlman School, Class Al (Blocks I-IV). Curriculum Document C-222-2010. Chief of Naval Technical Training, NAS Memphis, TN. 10 June, 1976 Revision.

Connolly, D.W. Voice Data Entry in Air Traffic Control. Technical Report FAA-NA-79-20, National Aviation Facilities Experimental Center, Atlantic City, NJ August, 1979.

Cronbach, L.J. and Snow, R.E. Aptitudes and Instructional Methods. Irvington Publishers, Inc., New York, 1977.

Dixon, N.R. and Martin, T.B. (Eds.) Automatic Speech and Speech Recognition, New York, IEEE Press, 1979.

Feuge, R.L., Charles, J.P. and Miller, R. Feasibility of Automated Adaptive GCA Controller Training System. Technical Report NAVTRAEQUIPCEN 73-C-0079-1. Naval Training Equipment Center, Orlando, FL. April, 1974.

Grady, M.W. and Hicklin, M. Use of Computer Speech Understanding in Training: A Demonstration Training System for the Ground Controlled Approach Controller. Technical Report NAVTRAEQUIPCEN 74-C-0048-1. Naval Training Equipment Center, Orlando, FL. December, 1976.

Grady, M.W., Hicklin, M.B., and Porter, J.E. AST in the 80's: New Systems, New Payoffs. Paper presented at the "Symposium on Voice Interactive Systems: Applications and Payoffs." Dallas, TX. May 13-15, 1980.

Harris, S. Proceedings of the Voice Interactive Systems: Applications and Payoffs Conference Dallas, TX, 13-15 May 1980, in press.

Hicklin, M., Nowell, L.H. and Petersen, R. Ground Controlled Approach Controller System Training/Functional Design Report. Technical Report NAVTRAEQUIPCEN 77-C-0162-1. Naval Training Equipment Center, Orlando, FL. December, 1978.

Hicklin, et al. Ground Controlled Approach Controller Training System (GCA-CTS) Student Guide. Technical Report NAVTRAEQUIPCEN 77-C-0162-4. Naval Training Equipment Center, Orlando, FL. 23 August, 1979a.

Hicklin, et al. Ground Controlled Approach Controller Training System (GCA-CTS) Instructor Guide. Technical Report NAVTRAEQUIPCEN 77-C-0162-5. Naval Training Equipment Center, Orlando, FL. 23 August 1979b.

Hicklin, M., Barber, G.C., Bollenbacher, J., Grady, M.W., Harry, D., Meyn, C., and Selemon, G.K. Ground Controlled Approach Controller Training System, Final Technical Report. Technical Report NAVTRAEQUIPCEN 77-C-0162-6. Naval Training Equipment Center, Orlando, FL. 23 February 1980.

Joplin, L. Interactive Logs For Automated Speech Technology. Technical Report NAVTRAEQUIPCEN 79-C-0066-1. Naval Training Equipment Center, Orlando, FL. 9 May 1980.

Lea, W. Critical Issues in Airborne Applications of Speech Recognition, Final Task Report submitted to NAVAIRDEVCEN by Speech Communications Research Laboratory, Los Angeles, CA, 1979.

Lea, W.A. (Ed.) Trends in Speech Recognition. Prentice-Hall, Inc., Englewood Cliffs, NJ. 1980.

Logicon, Inc. Ground Controlled Approach Controller Training System Technical Description. Logicon, Inc., Tactical and Training Systems Division, San Diego, CA. September, 1979.

Mecherifkoff, M., and Mackie, R.R. Attitudinal Factors in the Acceptance of Innovations in the Navy. Technical Report 784-1. Human Factors Research, Inc., Goleta, CA, June, 1970.

Phares, E.J. Locus of Control in Personality. Morristown, N.J: General Learning Press, 1976.

Powers, W.T. Behavior: The Control of Perception. Aldine Publishing Company, Chicago, IL. 1973.

Reed, V.J. Factors Affecting Attrition in AC(A1) School at NATTC, Memphis. Evaluation Branch Attrition Report. Chief of Naval Technical Training, NAS Memphis, TN. November, 1977.

APPENDIX A
ACRONYMS

AIC	Air Intercept Controller
APE	Aircraft/Pilot/Environmental model
ASR	Air surveillance radar
AST	Automated speech technology
ASVAB	Armed Services Vocational Aptitude Battery
ATC	Air traffic control
CAI	Computer assisted instruction
CATCC	Carrier Air Traffic Control Center
CRT	Cathrode ray tube
CSR	Computer speech recognition
GCA	Ground controlled approach
GCA-CTS	Ground Controlled Approach-Controller Training System
ILS	Instrument landing system
IWR	Isolated word recognition
LSO	Landing Signal Officer
NAS	Naval Air Station
NATTC	Naval Air Technical Training Center
NAVAID	Navigational aid
NAVTRAEQUIPCEN	Naval Training Equipment Center
OJT	On-the-job training

APPENDIX A
(CONTINUED)

PAR	Precision approach radar
PARTS	Precision Approach Radar Training System
P-run	Performance run
RATCF	Radar air traffic control facility
SUS	Speech understanding system
TCC	Target control console
TEE	Training effectiveness evaluation
TOT	Transfer of training
TRA	Transmission recognition accuracy

Naval Training Equipment Center
Orlando, FL 32813 65

Defense Technical Information Center
Cameron Station
Alexandria, VA 22314 12

All other addressees receive one copy

OUSDR&E (R&AT) (E&LS)
CDR Paul R. Chatelier
Washington, DC 20301

Commander
Naval Air Systems Command
Washington, DC 20361

Mr. Walter Primas
Chief of Naval Operations
OP-39T
Washington, DC 20350

Chief of Naval Operations
OP-593B
Washington, DC 20350

Chief of Naval Operations
OP-513
Washington, DC 20350

Chief of Naval Operations
OP-987H
Attn: Dr. R. G. Smith
Washington, DC 20350

Deputy Chief of Naval Operations (MPT)
Attn: OP-102
Arlington Annex
Washington, DC 20350

Mr. Ernest Poor
Naval Air Systems Command
AIR 413B
Room 338
Washington, DC 20361

Commander
Naval Air Systems Command
AIR 340F
Washington, DC 20361

Chief of Naval Research
Code 458
800 N. Quincy Street
Arlington, VA 22217

Dr. Marshall J. Farr
Director, Personnel and Training Research Program
Office of Naval Research
Arlington, VA 22217

Chief of Naval Air Training
Attn: Code 333
NAS
Corpus Christi, TX 78419

T. Weiner
Naval Air Development Center
Code 4043
Warminster, PA 18974

Dr. Sam Schiflett
Naval Air Test Center
SY 721
Patuxent River, MD 20670

CDR Richard S. Gibson
Bureau of Medicine and Surgery
Head, Aerospace Psychology Branch
Code 3013
Washington, DC 20372

Chief of Naval Material
MAT 08D2
CP5, Room 678
Attn: Arnold I. Rubinstein
Washington, DC 20360

Mr. Harry A. Whitted
Code 8235
Naval Ocean Systems Center
271 Catalina Boulevard
San Diego, CA 92152

Dr. Richard J. Schifflec
US Air Force
ASD/ENECH
Wright-Patterson AFB
Dayton, OH 45433

Major Neal Morgan (LGY)
Air Force Logistics Management Center
Bldg. 205
Gunter AFS, AL 36114

Mr. Harold A. Kottmann
ASD/YWE
Wright Patterson AFB, OH 45433

Mr. Thomas J. Moore
AFAMRL/BBA
Wright-Patterson AFB
Dayton, OH 45433

US Air Force Human Resource Lab
AFHRL-FT
Flying Training Division
Attn: Dr. Edwards
Williams AFB, AZ 85224

US Air Force Human Resources Lab
TSZ
Brooks AFB, TX 78235

Dr. J. D. Fletcher
Defense Adv. Research Projects Agency
(CTO)
1400 Wilson Boulevard
Arlington, VA 22209

AFHRL/PE
Brooks AFB, TX 78235

Mr. Eric Werkowitz
AFFDL/FGR
Wright-Patterson AFB, OH 45433

Chief
ARI Field Unit
P. O. Box 476
Ft. Rucker, AL 36362

Commanding Officer
US Army Security Agency Training Center
Library
Ft. Devens, MA 01433

ASD/AXA
Attn: N. R. Vivians
Wright-Patterson AFB, OH 45433

LTCOL Ray Kreulen
USAAVNO, DCAT-DT (ATZO-T-TM)
P. O. Box 365
Ft. Rucker, AL 36362

USAHEL/USAAVNC
Attn: DRXHE-FR
P. O. Box 476
Ft. Rucker, AL 36362

Director
US Army Research and Technology Lab
DAVDL-AS (Attn: Dr. R. S. Dunn)
207-5 Ames Research Center
Moffett Field, CA 94035

Dr. Donald W. Connolly
Research Psychologist
Federal Aviation Administration
FAA/NAFEC ANA-230 Bldg. 3
Atlantic City, NJ 08405

Dr. Jesse Orlansky
Institute for Defense Analyses
Science and Technology Division
400 Army-Navy Drive
Arlington, VA 22202

National Aviation Facilities
Experimental Center
Library
Atlantic City, NJ 08405

Mr. Clayton R. Coler
Research Scientist
NASA Ames Research Center
Mail Stop 239-2
Moffett Field, CA 94035

Commander
Naval Air Force
US Pacific Fleet (OOIE-31E)
NAS North Island
San Diego, CA 92135

Dr. Henry DeHaan
US Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. John F. Boehm
Director, National Security Agency
9800 Savage Road
Attn: R-542
Ft. George G. Meade, MD 20755

Scientific Advisor
Headquarters US Marine Corps
Washington, DC 20360

Commander
Naval Air Development Center
Attn: Code 6022
Warminster, PA 18974

Chief of Naval Technical Training
Code 0161
NAS Memphis (75)
Millington, TN 38054

Mr. Gary Poock
Naval PG School
Code 55PK
Monterey, CA 93940

Mr. John Martins, Jr.
Project Engineer
Naval Underwater Systems Center
New London Laboratory MC 315
New London, CT 06320

Commanding Officer
Rome Air Development Center
Library (TSLD)
Griffiss AFB, NY 13440

Commanding Officer
Fleet Anti-Submarine Warfare Training Center, Pacific
Attn: Code 001
San Diego, CA 92147

Chief of Naval Reserve
Code S-33
New Orleans, LA 70146

Mr. John Silva
Naval Ocean Systems Center
Code 823
San Diego, CA 92152

Navy Personnel Research Air Development Center
Attn: M. McDowell
Library, Code P201L
San Diego, CA 92152

Commanding Officer
Fleet Combat Training Center, Pacific
Code 04A
San Diego, CA 92147

Commanding Officer
Naval Aerospace Medical Research Lab
Code L5
Department of Psychology
Pensacola, FL 32512

Commander
Naval Air Force
US Pacific Fleet (Code 342)
NAS North Island
San Diego, CA 92135

Mr. Robert H. Wright
Engineering Research Psychologist
Army Aeromechanics Lab, R&T Labs
MS 239-2, MVSRD, Ames Research Center
Moffett Field, CA 94035

Air Force Institute Technology Library
Wright-Patterson AFB, OH 45433

Headquarters
Air Training Command XPTI
Attn: Mr. Goldman
Randolph AFB, TX 78148

ASD/ENETC
Mr. R. G. Cameron
Wright-Patterson AFB, OH 45433